Optimizing multilayered diffractive optical elements for compressive sensing in imaging systems

Tomoya Nakamura[a] and Mohamad Ammar Alsherfawi Aljazaerly[a]

[a]SANKEN, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, Japan

ABSTRACT

This study explored the optimization of multi-layered diffractive optical elements (DOEs) for compressive sensing in imaging systems. We investigated how increasing the number of DOE layers with optimization affected the coherence value of the system's transmission matrix, a key performance indicator for image reconstruction in compressive sensing. Through wave-optics simulations, we demonstrated that coherence decreased monotonically up to 8 layers before increasing, suggesting an optimal layer count exists. Our findings indicated that optimized multi-layer DOEs can significantly enhance the performance of compressive sensing applications, potentially improving accuracy in snapshot super-resolution and multi-dimensional imaging.

Keywords: Diffractive optical element, Compressive sensing, Optimization

1. INTRODUCTION

Computational coded imaging is a hybrid approach that combines optically coded measurement and computational decoding to achieve image aquisition. Among such approaches, pupil phase coding by using a diffractive optical element (DOE), which realizes point spread function (PSF) engineering, has been widely studied for its potential to improve the performance of imaging systems. For example, past studies on DOE-based PSF engineering enabled snapshot digital super-resolution,[1,2] three-dimensional imaging,[3] hyperspectral imaging,[4] depth-of-field extension,[5–7] lensless imaging,[8,9] and so on.

An important aspect of DOE-based PSF engineering is the ability to integrate the compressive-sensing (CS) framework with an imaging system.[10,11] CS framework in imaging is a method for reconstructing a higher-dimensional and/or larger-resolution image than a captured data. To achieve this reconstruction by continuous optimization method, particularly ℓ_1-norm minimization method, the transmission matrix,[12,13] which is a matrix to describe the linear forward process of image sensing, must have a low coherence value.[14] The transmission matrix $\mathbf{H}$ is typically represented as a set of the column vector $\mathbf{a}_i$ which represents an impulse responce of a space-variant point-source input as:

$$\mathbf{H} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N], \tag{1}$$

where N is the number of pixels in the image. Note that the $\mathbf{a}_c$ where c is the center of the image is defined as the PSF if the impulse response has the space-invariant characteristics.[12] The coherence value μ of the transmission matrix is defined as the maximum value of the column-vector correlation as:

$$\mu = \max_{i,j} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2}, \tag{2}$$

where $\mathbf{a}_i$ and $\mathbf{a}_j$ are the i-th and j-th column vectors of the transmission matrix, respectively. If the low-coherence value is satisfied in the matrix by the appropreate optical design, the higher-dimensional and/or larger-resolution image $\hat{\mathbf{x}}$ can be numerically reconstructed from given measurement data $\mathbf{y}$ by the error minimization method with the sparsity constraint as:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{H}\mathbf{y} - \mathbf{x}\|_2^2 + \tau\Psi(\mathbf{x}), \tag{3}$$

Further author information: (Send correspondence to T.N.)
T.N.: E-mail: nakamura.tomoya.sanken@osaka-u.ac.jp, Telephone: +81 6 6879 8422

Practical Holography XXXIX: Displays, Materials, and Applications, edited by Pierre-Alexandre J. Blanche, Hiroshi Yoshikawa, Proc. of SPIE Vol. 13390, 133900G · © 2025 SPIE · 0277-786X · doi: 10.1117/12.3040582

where τ is a regularization parameter, and $\Psi(\mathbf{x})$ is a regularization function. This minimization problem can be solved by using the iterative algorithm such as ADMM[15] or TwIST[16] algorithms.

As mentioned above, the design of DOE for implementing the low-coherence transmission matrix is a key to achieve the high performance of CS framework in imaging. In previous studies, the random diffuser or random mask is known as a good design of DOE for the low-coherence matrix.[8,17] Recently, the optimization of the DOE design has been studied to improve the performance of CS-based imaging.[4,18] However, the most of previous studies have focused on the optimization of a single-layer DOE. This strategy can flexibly design the single PSF, but the freedom of design of the *matrix* is limited by the shift-invariance of the PSF.

To overcome this limitation, we propose the use of multi-layered DOEs, which can provide additional degrees of freedom in the design of the transmission matrix. In our previous study on computational lensless imaging, we demonstrated that the multi-layerd random coded apertures can achieve the lower condintion-number transmission matrix than the single coded aperture.[19] In this study, we extend the concept of this *matrix-engineering by multi-layerd coded optics* to lense- and DOE-based CS imaging, and invesitigate the effect of optimization of the coded optics for the multi-layered system. In this study, we also investigate the effect of the number of DOE layers on the coherence value through simulation results.

2. METHOD

2.1 Forward model of imaging with multi-layer DOEs

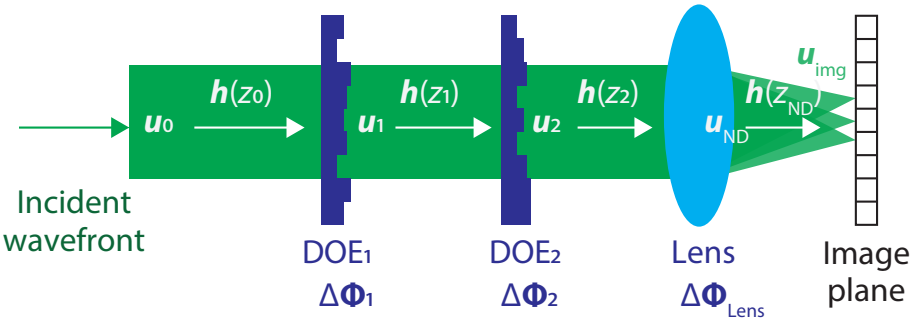

Figure 1. Schematic diagram of an imaging system with multi-layered DOEs.

Figure 1 shows the schematic of the imaging system with multi-layered DOEs. A vector of a complex wave field $\mathbf{u}$ can be represented by the following equation:

$$\mathbf{u} = \mathbf{A}\exp\left(jk\mathbf{\Phi}\right), \tag{4}$$

where $\mathbf{A}$ is a vector of the amplitude distribution of the wave field, k is the wavenumber, and $\mathbf{\Phi}$ is a vector of the phase distribution of the wave field. The phase modulation $\Delta\mathbf{\Phi}$ by the i-th DOE can be described by the following equation:

$$\Delta\mathbf{\Phi}_i = -n\mathbf{l}_i, \tag{5}$$

where n is the refractive index of the DOE, and $\mathbf{l}_i$ is a vector of the thickness of the i-th DOE. Under the Fresnel approximation, the propagation of the wave field to distance z from i-th DOE plane to $i+1$-th DOE plane can be simply described by the following equation:[12]

$$\mathbf{u}_{i+1} = \mathbf{h}(z_i) * \mathbf{u}_i, \tag{6}$$

where $\mathbf{h}(z_i)$ is a vector of the Fresnel kernel with the propagation distance z_i. After the modulations by the multi-layered DOE system, the wave field at the output plane is converted to the convering wave for focusing near the image plane. The phase modulation by the lens can be described by the following equation:

$$\Delta\mathbf{\Phi}_{\mathrm{Lens}} = n\mathbf{l}_{\mathrm{lens}} = \frac{x^2+y^2}{2f}, \tag{7}$$

where x and y are the spatial coordinates, and f is the focal length of the lens. By using above equations, the forward model of the imaging system with multi-layered DOEs can be described by the algorithm 1. In the algorithm, N_{D} is the total number of the DOE, $\odot$ denotes the Hadamard product, $*$ denotes the 2D convolution operation, and $\cdot^*$ denotes phase conjugate. Note that the $\mathbf{u}_0$ indicates the incident wave field to the DOE system, and the $\mathbf{I}_{\mathrm{img}}$ is a vector of the intensity distribution of the image. In the algorithm, several operations are cascaded, however the whole process can be optimized by gradient method because every operation is differentiable. To simulate spatial imaging at far field, we assumed tilted planer waves as the incident wave fields.

Algorithm 1 Forward Wave-Propagation Algorithm

for $i = 0$ to $N_{\mathrm{D}} - 1$ **do** ▷ Loop over each propagation layer
  if $i > 0$ **then**
    $\mathbf{u}_i \leftarrow \mathbf{u}_i \odot \exp(jk\Delta\mathbf{\Phi}_i)$ ▷ Phase modulation by a DOE to $\mathbf{u}_i$
  end if
  $\mathbf{u}_{i+1} \leftarrow \mathbf{h}(z_i) * \mathbf{u}_i$ ▷ Propagation of a wavefront
end for
$\mathbf{u}_{N_{\mathrm{D}}} \leftarrow \mathbf{u}_{N_{\mathrm{D}}} \odot \exp(jk\Delta\mathbf{\Phi}_{\mathrm{Lens}})$ ▷ Phase modulation by a lens
$\mathbf{u}_{\mathrm{img}} \leftarrow \mathbf{h}(z_{N_{\mathrm{D}}}) * \mathbf{u}_{N_{\mathrm{D}}}$ ▷ Propagation to the image plane
$\mathbf{I}_{\mathrm{img}} \leftarrow \mathbf{u}_{\mathrm{img}} \odot \mathbf{u}^*_{\mathrm{img}}$ ▷ Calculate intensity at the image plane

2.2 Optimization

To optimize DOEs, the $[\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_N]$ are set to the optimization variables, and they are optimized to minimize the predefined loss function that evaluates the coherence of the transmission matrix. The most straightfoward design of the loss function is just the summation of column-vector correlation of the transmission matrix as follows:

$$\mathcal{L}_{\mathrm{coherence}} = \sum_{i,j} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2}, \tag{8}$$

where $\mathbf{a}_i$ and $\mathbf{a}_j$ are the i-th and j-th column vectors of the transmission matrix, respectively. Here, the max-selection operation in Eq. (2) has been replaced with a summation operation to ensure the differentiability of the computation. This design of loss function is simple and reasonable; however, the size of the transmission matrix for spatial imaging is typically so huge that the computation of the Eq. (8) is impractical. As a substitute, we focused on the PSF $\mathbf{a}_c$ as a representive impulse-response vector, and we define the alternative loss function using the PSF as:

$$\mathcal{L}_{\mathrm{ACF-MTF}} = \mathrm{ACF}(\mathbf{a}_c) + \gamma \frac{1}{\mathrm{MTF}(\mathbf{a}_c)}, \tag{9}$$

where $\mathrm{ACF}(\cdot)$ is the sum of 2D autocorrelation function of the PSF, $\mathrm{MTF}(\cdot)$ is the sum of modulation transfer function calculated by the PSF, and γ is a parameter to balance the two terms. The ACF evaluates the randomness of the PSF, and the MTF evaluates the continuity of the abundance of high-frequency components of the PSF. Therefore, both the ACF and reciprocal MTF term in the loss function and minimization of them contributes to the suppression of the coherence value of the transmission matrix. The optimization problem that we solved in simulation is defined as:

$$\min_{[\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_{N\mathrm{D}}]} \mathcal{L}_{\mathrm{ACF-MTF}}(\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_{N\mathrm{D}}). \tag{10}$$

3. SIMULATIONS

We performed wave-optics simulations to investigate the effect of the number of DOE layers on the coherence value of the transmission matrix with the optimization of DOEs. The DOE was designed with a size of 2048×2048

pixels with the pitch of 1.0 μm. Wavelength was set to 630 nm, and the focal length of the lens was set to 2.0 mm. The phase modulation map (i.e. thickness) of each DOE was optimized with the loss function of Eq. (9). γ in the loss function was set to 4. The optimization was implemented by using PyTorch and Adam optimizer.[20] We performed the optimization operation with 10,000 epochs, and learning rate was set to 1e-07.

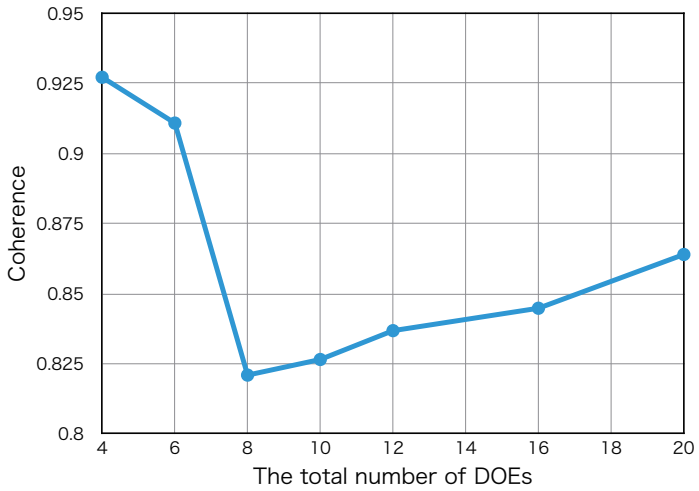

Figure 2. Coherence values of the transmission matrix with increasing the total number of DOEs.

Figure 2 shows the coherence value of the transmission matrix as a function of the number of DOE layers. Note that the coherence value for each layer count was calculated using the DOE optimized under each respective condition. As shonw in the figure, the coherence value decreased monotonically up to 8 DOEs before increasing. This result suggests that an optimal number of layers exists for minimizing the coherence value. The coherence value of the transmission matrix was reduced to 0.821 with optimized 8-DOE system, which was significantly lower than the value of 0.997 with a single DOE system.

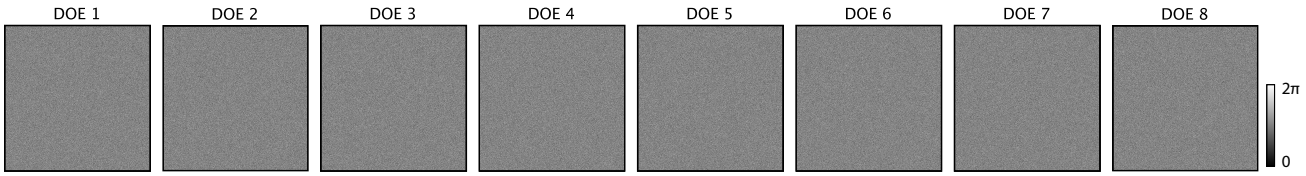

Figure 3. Phase-modulation maps of optimized 8-layer DOEs. The intensity values represent the phase delay amounts implemented by the DOE element.

Figure 3 shows the phase-modulation maps of the optimized 8-layer DOEs. The phase modulation map of the optimized DOE was effective in suppressing the coherence value, but it appeared to resemble a seemingly random pattern. This seemingly random yet designed phase modulation is thought to have structured the PSF randomly, resulting in output light with low correlation in response to the shift of the input point source. Consequently, it implemented a matrix with a lower coherence value.

4. CONCLUSIONS

In this study, we investigated whether an optical design method that involves multilayer placement and optimization of DOEs in a pupil-phase-modulation-based imaging optical system is effective in improving the performance of compressive sensing. The analysis results from wave optics simulations revealed that the coherence value of the transmission matrix, which is one of the indicators of compressive sensing performance, can be improved through the multilayering and optimization of DOEs. Furthermore, it was revealed that there is an optimal

number of layers in the design of multilayer DOE optical systems. Additionally, it was visualized that the DOE's height maps output from the optimization process exhibit an apparently random structure.

Our findings indicated that the optical design involving optimized multi-layer DOEs can significantly enhance the performance of CS-based imaging, potentially improving accuracy in CS-based snapshot super-resolution and multi-dimensional imaging. In future work, further performance analysis of multilayer DOE optical systems and proof-of-concept demonstrations based on physical implementation will be conducted.

ACKNOWLEDGMENTS

This work was supported by the JSPS KAKENHI (22H01992) and JST FOREST (JPMJFR206K).

REFERENCES

[1] Ashok, A. and Neifeld, M. A., "Pseudorandom phase masks for superresolution imaging from subpixel shifting," *Applied Optics* **46**, 2256–2268 (apr 2007).

[2] Kawachi, H., Nakamura, T., Iwata, K., Makihara, Y., and Yagi, Y., "Snapshot super-resolution indirect time-of-flight camera using a grating-based subpixel encoder and depth-regularizing compressive reconstruction," *Optics Continuum* **2**, 1368–1383 (jun 2023).

[3] Pavani, S. R. P., Thompson, M. A., Biteen, J. S., Lord, S. J., Liu, N., Twieg, R. J., Piestun, R., and Moerner, W. E., "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proceedings of the National Academy of Sciences* **106**, 2995–2999 (mar 2009).

[4] Baek, S.-H., Ikoma, H., Jeon, D. S., Li, Y., Heidrich, W., Wetzstein, G., and Kim, M. H., "Single-shot Hyperspectral-Depth Imaging with Learned Diffractive Optics," in [*2021 IEEE/CVF International Conference on Computer Vision (ICCV)*], 2631–2640, IEEE (oct 2021).

[5] Dowski, E. R. and Cathey, W. T., "Extended depth of field through wave-front coding," *Applied Optics* **34**, 1859–1866 (apr 1995).

[6] Cossairt, O. S., Zhou, C., and Nayar, S. K., "Diffusion Coded Photography for Extended Depth of Field," *ACM Transactions on Graphics* **29**(4), 31 (2010).

[7] Silva Neto, J. R. C. S. A. V., Nakamura, T., Makihara, Y., and Yagi, Y., "Extended Depth-of-Field Lensless Imaging Using an Optimized Radial Mask," *IEEE Transactions on Computational Imaging* **9**(3), 857–868 (2023).

[8] Antipa, N., Kuo, G., Reinhard, H., Mildenhall, B., Bostan, E., Ng, R., and Waller, L., "DiffuserCam: lensless single-exposure 3D imaging," *Optica* **5**(1), 1–9 (2018).

[9] Nakamura, T., Watanabe, T., Igarashi, S., Chen, X., Tajima, K., Yamaguchi, K., Shimano, T., and Yamaguchi, M., "Superresolved image reconstruction in FZA lensless camera by color-channel synthesis," *Optics Express* **28**, 39137–39155 (dec 2020).

[10] Donoho, D., "Compressed sensing," *IEEE Transactions on Information Theory* **52**, 1289–1306 (apr 2006).

[11] Stern, A., [*Optical Compressive Imaging*], CRC Press (2016).

[12] Goodman, J. W., [*Introduction to Fourier Optics*], McGraw-Hill (1996).

[13] Popoff, S. M., Lerosey, G., Carminati, R., Fink, M., Boccara, A. C., and Gigan, S., "Measuring the Transmission Matrix in Optics : An Approach to the Study and Control of Light Propagation in Disordered Media," *Phys. Rev. Lett.* **104**(100601), 1–4 (2010).

[14] Candès, E. and Romberg, J., "Sparsity and incoherence in compressive sampling," *Inverse Problems* **23**(3), 969–985 (2007).

[15] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning* **3**(1), 1–122 (2010).

[16] Bioucas-Dias, J. M. and Figueiredo, M. A. T., "A New TwIST : Two-Step Iterative Shrinkage / Thresholding Algorithms for Image Restoration," *IEEE Transactions on Image Processing* **16**(12), 2992–3004 (2007).

[17] Nakamura, T., Kagawa, K., Torashima, S., and Yamaguchi, M., "Super Field-of-View Lensless Camera by Coded Image Sensors," *Sensors* **19**(6), 1329 (2019).

[18] Sitzmann, V., Diamond, S., Peng, Y., Dun, X., Boyd, S., Heidrich, W., Heide, F., and Wetzstein, G., "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," *ACM Transactions on Graphics* **37**, 1–13 (aug 2018).

[19] Nakamura, T., Kato, R., Iwata, K., Makihara, Y., and Yagi, Y., "Multilayer lensless camera for improving the condition number," *Applied Optics* **63**, G9–G17 (oct 2024).

[20] Kingma, D. P. and Ba, J., "Adam: A Method for Stochastic Optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* , 1–15 (dec 2014).