

Noise-reduction algorithms for optimization-based imaging through multi-mode fiber

Ruo Yu Gu, Reza Nasiri Mahalati, and Joseph M. Kahn*

E. L. Ginzton Laboratory and Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA

**jmk@ee.stanford.edu*

Abstract: Three modifications are shown to improve resolution and reduce noise amplification in endoscopic imaging through multi-mode fiber using optimization-based reconstruction (OBR). First, random sampling patterns are replaced by sampling patterns designed to have more nearly equal singular values. Second, the OBR algorithm uses a point-spread function based on the estimated spatial frequency spectrum of the object. Third, the OBR algorithm gives less weight to modes having smaller singular values. In simulations for a step-index fiber supporting 522 spatial modes, the modifications yield a 20% reduction in image error (l^2 norm) in the noiseless case, and a further 33% reduction in image error at a 22-dB shot noise-limited SNR as compared to the original method using random sampling patterns and OBR.

©2014 Optical Society of America

OCIS codes: (060.2350) Fiber optics imaging; (100.3010) Image reconstruction techniques; (110.4280) Noise in imaging systems; (170.2150) Endoscopic imaging.

References and links

1. A. Yariv, "Three-dimensional pictorial transmission in optical fibers," *Appl. Phys. Lett.* **28**(2), 88 (1976).
2. B. Fischer and S. Sternklar, "Image transmission and interferometry with multimode fibers using self-pumped phase conjugation," *Appl. Phys. Lett.* **46**(2), 113 (1985).
3. R. N. Mahalati, R. Y. Gu, and J. M. Kahn, "Resolution limits for imaging through multi-mode fiber," *Opt. Express* **21**(2), 1656–1668 (2013).
4. I. N. Papadopoulos, S. Farahi, and C. Moser, "High resolution, lenseless endoscope based on digital scanning through a multimode optical fiber," *Opt. Express* **4**(2), 17598–17603 (2013).
5. T. Čížmár and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," *Nat. Commun.* **3**, 1027 (2012).
6. S. Bianchi and R. Di Leonardo, "A multi-mode fiber probe for holographic micromanipulation and microscopy," *Lab Chip* **12**(3), 635–639 (2012).
7. Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber," *Phys. Rev. Lett.* **109**(20), 203901 (2012).
8. S. Farahi, D. Ziegler, I. N. Papadopoulos, D. Psaltis, and C. Moser, "Dynamic bending compensation while focusing through a multimode fiber," *Opt. Express* **21**(19), 22504–22514 (2013).
9. A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, and R. Piestun, "Real-time resilient focusing through a bending multimode fiber," *Opt. Express* **21**(10), 12881–12887 (2013).
10. I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "Increasing the imaging capabilities of multimode fibers by exploiting the properties of highly scattering media," *Opt. Lett.* **38**(15), 2776–2778 (2013).
11. S. Bianchi, V. P. Rajamanickam, L. Ferrara, E. Di Fabrizio, C. Liberale, and R. Di Leonardo, "Focusing and imaging with increased numerical apertures through multimode fibers with micro-fabricated optics," *Opt. Lett.* **38**(23), 4935–4938 (2013).
12. Y. Choi, C. Yoon, M. Kim, J. Yang, and W. Choi, "Disorder-mediated enhancement of fiber numerical aperture," *Opt. Lett.* **38**(13), 2253–2255 (2013).
13. B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. M. Cheung, and M. J. Schnitzer, "Fiber-optic fluorescence imaging," *Nat. Methods* **2**(12), 941–950 (2005).
14. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
15. E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory* **52**(12), 5406–5425 (2006).
16. J. A. Buck, *Fundamentals of Optical Fibers*, 2nd ed. (John Wiley, 2004).
17. J. Mertz, *Introduction to Optical Microscopy* (Roberts and Company, 2010).

18. J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts and Company, 2004).
 19. G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed. (John Wiley, 2010), Chap. 3.
 20. A. J. Welch, M. J. C. Van Gemert, W. M. Star, and T. Optics, *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed. (Springer, 2011), Chap. 3.
 21. A. J. Welch, "The thermal response of laser irradiated tissue," *IEEE J. Quantum Electron.* **20**(12), 1471–1481 (1984).
 22. A. L. McKenzie, "Physics of thermal processes in laser-tissue interaction," *Phys. Med. Biol.* **35**(9), 1175–1210 (1990).
 23. S. W. Allison, G. T. Gillies, D. W. Magnuson, and T. S. Pagano, "Pulsed laser damage to optical fibers," *Appl. Opt.* **24**(19), 3140–3144 (1985).
 24. M. B. Shemirani, W. Mao, R. A. Panicker, and J. M. Kahn, "Principal modes in graded-index multimode fiber in presence of spatial- and polarization-mode coupling," *J. Lightwave Technol.* **27**(10), 1248–1261 (2009).
-

1. Introduction

The use of multi-mode fibers (MMFs) for imaging has long been of interest [1,2], and the past few years have seen particularly productive developments in the field of MMF endoscopic imaging [3–12]. An endoscope employing one MMF provides resolution superior to, and is much more compact than, current commercial fiber endoscopes, which may use either a bundle of single-mode fibers or one single-mode fiber with a scanning probe head [13]. Light signals propagating through a MMF are subject to random scattering by index imperfections, bends and other perturbations. All endoscopic MMF imaging methods, prior to image recording, employ some calibration method that requires access to both ends of the fiber. Since bending changes a fiber's propagation characteristics, all MMF endoscopes to date have been rigid, although compensation of bending has been investigated [8,9].

One endoscopic MMF imaging method [3] uses random intensity patterns to sample an object and employs an optimization-based image reconstruction (OBR) algorithm. This method at once enhances the resolution to four times the number of spatial modes propagating in the MMF and requires the least complicated calibration setup among all imaging methods. The OBR algorithm decomposes the set of random intensity sampling patterns into a set of intensity modes using a singular-value decomposition (SVD), and reconstructs the image as a linear combination of the intensity modes based on the powers reflected from the object when sampled by the random intensity patterns. This technique is similar to compressed sensing [14,15].

Despite the advantages of this imaging method, reconstructed images were found to be noisy [3], suggesting that noise amplification occurs in OBR. In this paper, we quantify the noise amplification occurring in OBR, and describe three methods to reduce it. First, we sample the object using a set of intensity patterns designed to have singular values more equalized than a set of random intensity patterns. Second, we modify the OBR algorithm to use a point-spread function (PSF) based on the estimated spatial frequency spectrum of the object. Third, we diminish the influence of intensity modes with smaller singular values on the reconstructed image. Using simulation, we compare images obtained using noise-reduced OBR to those obtained by localized spot sampling with localized reconstruction in the presence of noise, and subject to constraints on the number of sampling patterns and the power illuminating the object through the MMF.

The remainder of this paper is as follows. In Section 2, we describe the imaging system and a model for the electric fields illuminating the object and captured by the MMF, and briefly review the OBR algorithm described in [3]. In Section 3, we analyze the image noises obtained using OBR or localized sampling/reconstruction and quantify the noise amplification of OBR. In Section 4, we introduce three methods to improve the noise performance of OBR. In Section 5, we present simulated images obtained using OBR or localized sampling/reconstruction under the constraints mentioned above. We discuss the results in Section 6 and provide conclusions in Section 7.

2. Model for object sampling and image reconstruction

2.1 Imaging system and model for electric fields

The imaging system simulated here is identical to the experimental system in [3], except it uses a step-index MMF instead of the graded-index MMF used in [3]. As shown in Fig. 1, a 1550-nm laser beam illuminates a phase-only spatial-light modulator (SLM). The beam reflected from the SLM is coupled into a step-index MMF having 50- μm core diameter. The MMF is 1 m long, and its output facet is anti-reflection-coated with reflectivity $<1\%$. Calibration is performed by imaging intensity patterns appearing at the MMF output onto a camera with 256×256 pixels, as shown in Fig. 1(a). After calibration, an object to be imaged is placed at the MMF output, as shown in Fig. 1(b). A sequence of M different intensity patterns is displayed at the MMF output and is used to sample the object. When each intensity pattern is displayed, the total power reflected from the object and coupled back into the MMF is measured using a power meter. Then, using the sequence of M reflected power values, an image is reconstructed by OBR.

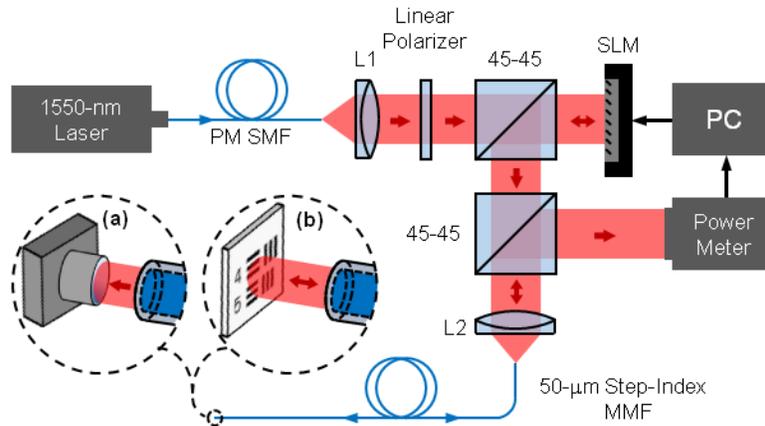


Fig. 1. Experimental setup for imaging through multi-mode fiber (MMF) [3]. (a) In calibration, a camera is placed at the MMF output to record intensity patterns appearing there. (b) In imaging, an object is placed at the MMF output and illuminated by a sequence of intensity patterns, and the reflected power values coupled back into the MMF are recorded.

A step-index MMF with core diameter d and numerical aperture NA supports $N \approx \frac{1}{4}V^2 = \frac{1}{4}(\pi d \text{NA}/\lambda)^2$ electric field modes per polarization at wavelength λ for large V [16]. For small NA, we can use the weakly guiding approximation and use analytically computed linearly polarized modes (LP) $E_m(\rho, \phi)$ as a basis. For MMFs with higher NAs, we instead use the numerically computed vector modes of the MMF, which deviate slightly from the analytically computed LP modes. The index l ranges from 0 to l_{\max} , while the index m ranges from 1 to m_{\max} , where both l_{\max} and m_{\max} depend on V [16]. Each $E_m(\rho, \phi)$ is two-fold degenerate for $l=0$, for the \hat{x} and \hat{y} polarizations; and fourfold degenerate for $l>0$, twice for the \hat{x} and \hat{y} polarizations and twice for a rotation of the mode about the \hat{z} axis by $\pi(2l)^{-1}$ radians (where \hat{z} is the direction of propagation). We assume that both the MMF and sample to be imaged have negligible birefringence; thus a single SLM provides full control of the intensity at the output of the MMF, and we can ignore polarization effects during imaging. For $l>1$, the complex electric field inside the MMF in one polarization is described in cylindrical coordinates by:

$$E_{lm}(\rho, \phi) = c_1 J_l(\kappa_{l,m} \rho) \sin(l\phi) \quad (\text{core}) \quad (1)$$

$$E_{lm}(\rho, \phi) = c_2 K_l(\gamma_{l,m} \rho) \sin(l\phi) \quad (\text{cladding})$$

where ρ and ϕ denote radius and azimuthal angle; $J_\nu(\rho)$ is the Bessel function of the first kind of order ν ; $K_\nu(\rho)$ is the modified Bessel function of the second kind of order ν ; $\kappa_{l,m}$ and $\gamma_{l,m}$ are constants that depend on NA, d and λ ; and c_1 and c_2 are real normalization constants. The expressions for the electric field for $l \leq 1$ are similar.

Let P1 denote the MMF output plane and P2 denote the object plane. At P1, the total forward-propagating electric field is a linear combination of the LP modes, and is described by:

$$E_{P1f}(x, y) = \sum_{l=0, m=0, q=0}^{l_{\max}, m_{\max}, q_{\max}} c_{lmq} E_{lmq}(x, y), \quad (2)$$

where c_{lmq} is a complex amplitude. A change of basis from cylindrical (ρ, ϕ) to Cartesian (x, y) co-ordinates has been made from Eq. (1), and the index q has been added to differentiate between the degeneracies of $E_{lm}(x, y)$.

Free-space propagation over a distance z is represented by $P_z[\]$, and is modeled here by Rayleigh-Sommerfeld diffraction [17]. At P2 the total forward-propagating electric field is thus:

$$E_{P2f}(x, y) = \sum_{l=0, m=0, q=0}^{l_{\max}, m_{\max}, q_{\max}} c_{lmq} P_z[E_{lmq}(x, y)], \quad (3)$$

At P2, a back-propagating electric field is generated by reflection from the object. We assume that the electric field reflects from a thin surface and reflection can be described by scalar diffraction theory [18], meaning that the reflected electric field is the product of the incident electric field and the complex reflectance of the object in each polarization on plane P2:

$$E_{P2b}(x, y) = r(x, y) E_{P2f}(x, y). \quad (4)$$

The power coupled back into the MMF is given by:

$$p = \sum_{l=0, m=0, q=0}^{l_{\max}, m_{\max}, q_{\max}} \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_{lmq}(x, y) P_{-z}[E_{P2b}(x, y)] dx dy \right|^2. \quad (5)$$

The power is assumed to be conserved in propagation through the MMF, and is measured by a power meter.

2.2 Optimization-based reconstruction algorithm

The relation between p and $r(x, y)$ in Eqs. (4) and (5) is not linear. Solving for $r(x, y)$ from p using Eq. (5) requires coherent detection of $E_{P2f}(x, y)$ and is in general computationally difficult for a MMF with the number of modes required for imaging. However, if we instead assume that the fraction of the reflected light recaptured by the MMF is independent of the object, and that the object $r(x, y)$ has equal reflectivity in both polarizations, Eq. (5) can be replaced by:

$$p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c |E_{p2f}(x, y) r(x, y)|^2 dx dy \quad (6)$$

$$p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\tilde{c} E_{p2f}(x, y)|^2 |r(x, y)|^2 dx dy \quad (7)$$

$$p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{I}_{p2f}(x, y) R(x, y) dx dy, \quad (8)$$

where c accounts for the fraction of the reflected light recaptured by the MMF for the sampling pattern $E_{p2f}(x, y)$ and any object $r(x, y)$; c can be computed using Eqs. (4) and (5) and assuming $r(x, y)$ is a perfect mirror, or measured experimentally. In Eq. (8), $\tilde{I}_{p2f}(x, y)$ is the intensity pattern at P2 measured by the camera during calibration and then multiplied by $|\tilde{c}|^2$. We note that in actuality c is not independent of $r(x, y)$, so the approximation used in Eq. (6) effectively introduces an error in p . The impact of this error will be seen in subsequent images, but the effect is small if the $E_{p2f}(x, y)$ used to sample the object are statistically similar (e.g. random patterns).

If we employ a sequence of M intensity patterns $\tilde{I}_{p2f}(x, y)$ during imaging, we obtain a corresponding $M \times 1$ vector of powers \mathbf{p} . If we discretize $\tilde{I}_{p2f}(x, y)$ and $R(x, y)$ to a grid of L points and reshape them into row vectors, we obtain the simple expression:

$$\mathbf{p} = \tilde{\mathbf{I}} \mathbf{r}, \quad (9)$$

where \mathbf{p} is an $M \times 1$ vector of all p , $\tilde{\mathbf{I}}$ is an $M \times L$ matrix in which each row represents an intensity sampling pattern, and \mathbf{r} is an $L \times 1$ vector representing the intensity reflectance of the object. This equation is linear, and the image \mathbf{w} can then be reconstructed using [3]:

$$\mathbf{w} = \mathbf{V} \tilde{\mathbf{D}}^{-1} \mathbf{U}^T \mathbf{p}, \quad (10)$$

where $\tilde{\mathbf{I}} = \mathbf{U} \mathbf{D} \mathbf{V}^T$ is the compact singular value decomposition of $\tilde{\mathbf{I}}$, and:

$$\tilde{D}_{i,j}^{-1} = \begin{cases} 1/D_{i,j} & \text{for } i = j \leq 4N \\ 0 & \text{otherwise} \end{cases}. \quad (11)$$

3. Noise of image reconstruction methods

In the previous section we assumed that the received power specified in Eq. (9) is noiseless. We now introduce noise. Given an apparatus, as the optical power is increased, the limiting noise source progresses from thermal to shot to intensity noise [19]. In the following section, we analyze all three regimes.

3.1 Noise of optimization-based image reconstruction

We assume that the calibration images $\tilde{\mathbf{I}}$ are noiseless. This assumption is valid if sufficient time is available for calibration using the camera to make noise negligible and there is no movement of the MMF during imaging. The noisy reconstructed image is then:

$$\mathbf{w} = \mathbf{V} \tilde{\mathbf{D}}^{-1} \mathbf{U}^T (\mathbf{p} + \mathbf{n}) \quad (12)$$

$$\mathbf{w} = \mathbf{V} \mathbf{V}^T \mathbf{r} + \mathbf{V}^T \tilde{\mathbf{D}}^{-1} \mathbf{U}^T \mathbf{n} \quad (13)$$

$$\tilde{\mathbf{w}} \equiv \mathbf{V}^T \tilde{\mathbf{D}}^{-1} \mathbf{U}^T \mathbf{n}, \quad (14)$$

where we have defined $\tilde{\mathbf{w}}$ as the image error due to noise. Here, \mathbf{n} is a vector of random noises, and depends on whether thermal, shot or intensity noise dominates. In the three cases, it is given by:

$$\mathbf{n} = \sigma_{\text{thermal}} \hat{\mathbf{n}} \text{ (thermal)} \quad (15)$$

$$\mathbf{n} = \sigma_{\text{shot}} \hat{\mathbf{n}} \circ \sqrt{\mathbf{p}} \text{ (shot)} \quad (16)$$

$$\mathbf{n} = \sigma_{\text{intensity}} \hat{\mathbf{n}} \circ \mathbf{p} \text{ (intensity)}, \quad (17)$$

where \circ is the Hadamard (elementwise) product and $\hat{\mathbf{n}}$ is a vector of i.i.d. Gaussian random variables with mean 0 and variance 1. The amount of noise amplification occurring in OBR can be quantified from the relation between the l^2 norm of $\tilde{\mathbf{w}}$, given by $E\left[\sum_{i=1}^L \tilde{w}_i^2\right]$, and \mathbf{n} .

For intensity noise, the relevant equations are:

$$\mathbf{p} = \mathbf{U} \mathbf{D} \mathbf{V}^T \mathbf{r} \quad (18)$$

$$\mathbf{n} = \sigma_{\text{intensity}} \hat{\mathbf{n}} \circ \mathbf{p} \quad (19)$$

$$\tilde{\mathbf{w}} = \mathbf{V} \tilde{\mathbf{D}}^{-1} \mathbf{U}^T \mathbf{n} \quad (20)$$

$$\tilde{\mathbf{n}} \equiv \mathbf{U}^T \mathbf{n}, \quad (21)$$

In Eq. (21), $\tilde{\mathbf{n}}$ is the component of the noise in the basis of intensity modes, which is useful in the following derivation. We model \mathbf{r} as a vector of i.i.d. uniformly distributed random variables between 0 and 1. Solving for $E\left[\sum_{i=1}^L \tilde{w}_i^2\right]$, we have:

$$E\left[\sum_{i=1}^L \tilde{w}_i^2\right] = E\left[\sum_{i=1}^{4N} (D_{ii}^{-1} \tilde{n}_i)^2\right] = \sum_{i=1}^{4N} D_{ii}^{-2} E[\tilde{n}_i^2] \approx \frac{1}{M} E\left[\sum_{i=1}^M \tilde{n}_i^2\right] \sum_{i=1}^{4N} D_{ii}^{-2} \quad (22)$$

$$E\left[\sum_{i=1}^M \tilde{n}_i^2\right] = E\left[\sum_{i=1}^M n_i^2\right] = \sum_{i=1}^M E[n_i^2] = \sum_{i=1}^M \sigma_{\text{intensity}}^2 \mu(p_i)^2 + \sigma_{\text{intensity}}^2 \sigma^2(p_i). \quad (23)$$

Throughout Eqs. (22) and (23), we have used the fact that unitary matrices preserve the l^2 norm. The approximation in Eq. (22) is accurate as long as all patterns in $\tilde{\mathbf{I}}$ have total powers within the same order of magnitude. Similarly, we find:

$$\mu(p_i) = E[\mathbf{i}_i^T \mathbf{r}] = E[r_i] \sum_{j=1}^L \tilde{I}_{ij} = \frac{1}{2} \sum_{j=1}^L \tilde{I}_{ij} \quad (24)$$

$$\sigma^2(p_i) = \frac{1}{M} \sigma^2(r_i) \sum_{i=1}^{4N} D_{ii}^2 = \frac{1}{12M} \sum_{i=1}^{4N} D_{ii}^2, \quad (25)$$

where \mathbf{i}_i^T is the i th row of $\tilde{\mathbf{I}}$. We note that $\sum_{i=1}^M \sigma_{\text{intensity}}^2 \mu(p_i)^2 \gg \sum_{i=1}^M \sigma_{\text{intensity}}^2 \sigma^2(p_i)$, and so discard that term in Eq. (23). Using Eqs. (22)–(25), we finally have:

$$E\left[\sum_{i=1}^L \tilde{w}_i^2\right] = \frac{1}{M} \left(\sum_{i=1}^{4N} D_{ii}^{-2}\right) \left[\sum_{i=1}^M \left(\frac{1}{2} \sum_{j=1}^L \tilde{I}_{ij}\right)^2\right] \sigma_{\text{intensity}}^2. \quad (26)$$

Equivalently, after dividing by the number of image pixels and taking the square root, the root mean square deviation (RMSD) of the image error is given by:

$$\sigma(\tilde{w}_i) = \sqrt{\frac{1}{L} \sum_{i=1}^{4N} D_{ii}^{-2}} \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\frac{1}{2} \sum_{j=1}^L \tilde{I}_{ij} \right)^2} \sigma_{\text{intensity}} \text{ (intensity noise)} \quad (27)$$

$$\sigma(\tilde{w}_i) = \sqrt{\frac{1}{L} \sum_{i=1}^{4N} D_{ii}^{-2}} \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\frac{1}{2} \sum_{j=1}^L \tilde{I}_{ij} \right)} \sigma_{\text{shot}} \text{ (shot noise)} \quad (28)$$

$$\sigma(\tilde{w}_i) = \sqrt{\frac{1}{L} \sum_{i=1}^{4N} D_{ii}^{-2}} \sigma_{\text{thermal}} \text{ (thermal noise)}, \quad (29)$$

where the expressions for shot and thermal noise are found in a similar manner. We see that for all three types of noise, $\sigma(\tilde{w}_i)$ is proportional to $\sqrt{\frac{1}{L} \sum_{i=1}^{4N} D_{ii}^{-2}}$, but with different multiplicative factors.

3.2 Noise of localized reconstruction

In this imaging method, the sampling patterns at the MMF output plane P2 are localized spots on a regular grid of points. As each sampling pattern illuminates the object, the power reflected from the object yields the intensity in the corresponding region of the reconstructed image, a process termed localized reconstruction [3]. For localized reconstruction using M spots, the corresponding relation between the noise \mathbf{n} , spot sampling patterns $\tilde{\mathbf{I}}$ and image \mathbf{w} is:

$$\mathbf{w} = \mathbf{T} \left[\mathbf{s} \circ (\tilde{\mathbf{I}} \mathbf{r} + \mathbf{n}) \right] \quad (30)$$

$$\tilde{\mathbf{w}} = \mathbf{T} (\mathbf{s} \circ \mathbf{n}), \quad (31)$$

where \mathbf{T} is a $L \times M$ matrix that implements sinc interpolation to upsample the image, each row of $\tilde{\mathbf{I}}$ represents samples of the smallest spot that can be formed at the corresponding grid point using the propagating modes of the MMF [3], and \mathbf{s} is a $M \times 1$ vector of the reciprocal of the power that couples back into the MMF when imaging a perfect mirror using $\tilde{\mathbf{I}}$. The noises \mathbf{n} are as given in Eqs. (15)–(17). The corresponding RMSDs of image error due to noise are:

$$\sigma(\tilde{w}_i) = \frac{1}{2} \sigma_{\text{intensity}} \text{ (intensity noise)} \quad (32)$$

$$\sigma(\tilde{w}_i) = \sqrt{\frac{1}{2M} \sum_{i=1}^M \left(\sum_{j=1}^P \tilde{I}_{ij} \right)^{-1}} \sigma_{\text{shot}} \text{ (shot noise)} \quad (33)$$

$$\sigma(\tilde{w}_i) = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\sum_{j=1}^P \tilde{I}_{ij} \right)^{-2}} \sigma_{\text{thermal}} \text{ (thermal noise)}, \quad (34)$$

where we have used the fact that sinc interpolation leaves $\sigma(\tilde{w}_i)$ unchanged. With these equations, we may define noise amplification as the ratio of the image error RMSDs for OBR, given by Eqs. (27)–(29), to the image error RMSDs using localized reconstruction, given by Eqs. (32)–(34).

4. Techniques to improve noise performance

4.1 Sampling using optimal patterns

The dependence of the image error $\sigma(\tilde{w}_i)$ on the singular values of $\tilde{\mathbf{I}}$ in Eqs. (27)–(29) suggests that noise amplification can be reduced by choosing a specific set of sampling patterns $\tilde{\mathbf{I}}$ with more equalized singular values. We heuristically expect a grid of spots

(identical to that used in localized sampling and reconstruction) to be a good set of sampling patterns: among all possible sampling patterns satisfying the physical constraint of positive intensity ($\tilde{I}_{ij} \geq 0 \forall i, j$), the overlap integral between different sampling patterns is minimized

for a grid of spots, which is a good characteristic for minimizing $\sqrt{\sum_{i=1}^{4N} D_{ii}^{-2}}$. We briefly note that forming specific sampling patterns is necessarily more difficult than forming random patterns, since light undergoes random mode coupling during MMF propagation. Conveniently, methods have been described to form the entire set of spot patterns with total time linearly proportional to the number of modes N [4–6].

The impacts of shot and thermal noise are inversely dependent on the integrated total power in the sampling patterns, so it is desirable to maximize the total power in these patterns. However, this power must be limited to prevent damage to tissue [20–22], which occurs at powers well below those that cause any damage to the MMF [23]. Tissue damage is typically caused by thermal effects, which depend on the irradiance incident on the tissue [22], so we constrain the maximum irradiance of each sampling pattern such that $\tilde{I}_{ij} < \tilde{I}_{\max} \forall i, j$ (here we use the terms irradiance and intensity interchangeably in referring to the power per unit area). This constraint is particularly important when MMFs with high NA are used, as they permit the total power to be concentrated in a high-irradiance spot that is much smaller than the MMF core. However, since thermal effects are not directly localized to the irradiated area, a large and uniform sampling pattern can cause more damage than a smaller one with the same irradiance. To account for this, we additionally constrain the total power of each sampling pattern such that $\sum_{j=1}^L \tilde{I}_{ij} < P_{\max} \forall i$.

We plot the noise amplification for OBR when imaging using random and spot sampling patterns in Fig. 2, for both constraints. The intensity patterns have been generated using the model of Section 2, and the noise amplification is calculated using Eqs. (27)–(29) and Eqs. (32)–(34) of Section 3.

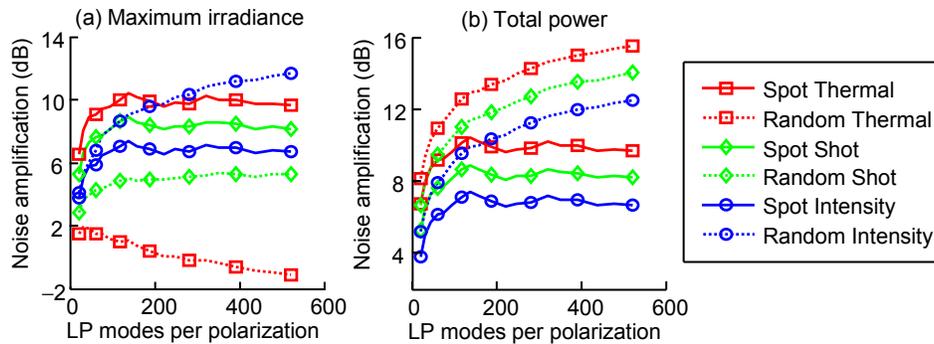


Fig. 2. Noise amplification for OBR when sampling using random (dashed) and spot (solid) intensity sampling patterns as a function of the number of LP modes per polarization under constraints on (a) maximum irradiance and (b) total power. Noise amplification values are obtained from Eqs. (27)–(29) and Eqs. (32)–(34). Intensity patterns are observed $5 \mu\text{m}$ from the MMF fiber tip. The MMF has a $50\text{-}\mu\text{m}$ core diameter and a varying NA. The number of sampling patterns is $16N$ (corresponding to $2 \times$ oversampling in the x - and y -dimensions for spots). Intensity patterns are sampled on a 416×416 grid.

Under the total power constraint shown in Fig. 2(b), when imaging with random patterns the noise amplification increases in proportion to the square root of the number of LP modes, while for imaging with spot patterns it remains constant. This applies to all types of noise. Thus imaging using spot patterns offers a clear improvement over imaging with random patterns under the total power constraint, with the drawback of a more difficult calibration of

the imaging apparatus. Conversely when operating under the maximum irradiance constraint of Fig. 2(a), spot patterns have less noise amplification than random patterns only for intensity noise. Thus imaging using random patterns is superior unless intensity noise dominates.

The results of Fig. 2 are for a specific MMF core diameter of 50 μm , but we expect the results to be similar for a constant number of modes independent of core diameter (with the distance to the object scaled according to the Rayleigh range of the MMF [3]). We note also that MMFs are commonly available with well over 500 modes, but it may not be possible to use all these modes for imaging due to time constraints; for example, for 500 modes and a grid of spots with an oversampling factor of $2 \times$ in the x - and y -dimensions, 8000 spots need to be sequentially formed during imaging to exploit all the modes.

4.2 Exploiting known properties of the object

In [3], the image reconstruction algorithm was chosen under the assumption that nothing is known about \mathbf{r} , the object to be imaged. However, at a minimum we do know that \mathbf{r} must be positive and have a limited spatial frequency bandwidth. If instead \mathbf{r} were i.i.d., it would have a uniform spatial frequency spectrum up to the spatial sampling rate of the camera used to record the intensity patterns. For realistic objects, the spatial frequency spectrum drops off well before this. With this motivation, we modify the OBR algorithm to exploit any information known about the spatial frequencies of the object. This requires a more general objective function than the one described in [3], which was:

$$\arg \min_{\mathbf{w}} \|\mathbf{p} - \tilde{\mathbf{I}}\mathbf{w}\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0}, \quad (35)$$

where \mathbf{p} , $\tilde{\mathbf{I}}$ and \mathbf{w} are defined in Section 2, \mathbf{V} is the matrix of right singular vectors defined in Section 2, and \mathbf{V}_2 is a submatrix of \mathbf{V} such that $\mathbf{V} = [\mathbf{V}_1 \quad \mathbf{V}_2]$, with \mathbf{V}_1 containing the right singular vectors that have singular values deemed to be significant (chosen here as $>1\%$ of the maximum singular value).

Instead of Eq. (35), we minimize the l^2 norm between the object intensity reflectance and the image:

$$\arg \min_{\mathbf{w}} \|\mathbf{w} - \mathbf{r}\|_2 \quad (36)$$

Note that we do not know \mathbf{r} . For consistency, we show below that Eq. (35) is equivalent to Eq. (36) with the constraint that $\mathbf{p} = \tilde{\mathbf{I}}\mathbf{w}$ when \mathbf{r} is i.i.d. with mean 0. For Eq. (35):

$$\begin{aligned} \arg \min_{\mathbf{w}} \|\mathbf{p} - \tilde{\mathbf{I}}\mathbf{w}\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0} &= \arg \min_{\mathbf{w}} \|\mathbf{U}\mathbf{D}\mathbf{V}^T \mathbf{r} - \mathbf{U}\mathbf{D}\mathbf{V}^T \mathbf{w}\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0} \\ &= \arg \min_{\mathbf{w}} \|\mathbf{D}\mathbf{V}^T (\mathbf{r} - \mathbf{w})\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0} \quad (37) \\ &= \arg \min_{\mathbf{w}} \|\mathbf{V}_1^T (\mathbf{w} - \mathbf{r})\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0} \end{aligned}$$

In the above, the final equality is valid because the minimum of the norm is 0. We have made the approximation that small singular values are zero. For Eq. (36):

$$\begin{aligned} \arg \min_{\mathbf{w}} \|\mathbf{w} - \mathbf{r}\|_2 \quad \text{s.t. } \mathbf{p} = \tilde{\mathbf{I}}\mathbf{w} &= \arg \min_{\mathbf{w}} \|\mathbf{V}^T (\mathbf{w} - \mathbf{r})\|_2 \quad \text{s.t. } \mathbf{V}_1^T \mathbf{r} = \mathbf{V}_1^T \mathbf{w} \\ &= \arg \min_{\mathbf{w}} \left(\|\mathbf{V}_1^T (\mathbf{w} - \mathbf{r})\|_2 + \|\mathbf{V}_2^T (\mathbf{w} - \mathbf{r})\|_2 \right) \quad (38) \\ &= \arg \min_{\mathbf{w}} \|\mathbf{V}_1^T (\mathbf{w} - \mathbf{r})\|_2 \quad \text{s.t. } \mathbf{V}_2^T \mathbf{w} = \mathbf{0} \end{aligned}$$

In the above, the final line is valid because \mathbf{r} is i.i.d. and equally likely to be negative or positive, so $\mathbf{V}_2^T \mathbf{w} = \mathbf{0}$ minimizes the l^2 norm. Equation (38) has the same result as Eq. (37) above.

We note that the continuous form of Eq. (36) allows a simplification if we assume that the point-spread function $q(x, y)$ of the image is spatially invariant. In that case:

$$\iint_{\text{Image}} [w(x, y) - r(x, y)]^2 dx dy \quad (39)$$

$$\iint_{\text{Image}} [r(x, y) * (q(x, y) - \delta(x, y))]^2 dx dy \quad (40)$$

$$\iint_{\text{Image}} [\tilde{r}(x, y)(\tilde{q}(x, y) - \tilde{\delta}(x, y))]^2 dx dy, \quad (41)$$

where \tilde{r} , \tilde{q} and $\tilde{\delta}$ are the Fourier transforms of the original functions and $*$ denotes convolution.

In minimizing Eq. (36), we restrict ourselves to solutions of the form:

$$\mathbf{w} = \tilde{\mathbf{V}} \tilde{\mathbf{D}}^{-1} \mathbf{U}^T \mathbf{p}, \quad (42)$$

where \mathbf{U}^T and $\tilde{\mathbf{D}}^{-1}$ are the same as in Eq. (10), and $\tilde{\mathbf{V}}$ remains to be determined. We chose Eq. (42) since it can be simplified to $\mathbf{w} = \tilde{\mathbf{V}} \mathbf{V}_1^T \mathbf{r}$, which allows us to interpret $\tilde{\mathbf{v}}_i^T \mathbf{V}_1^T$ as the point-spread function of the image at pixel i , where $\tilde{\mathbf{v}}_i^T$ is the i th row of $\tilde{\mathbf{V}}$. We know the objective function for the optimal PSF from Eq. (41), and so we use the following equation to solve for every $\tilde{\mathbf{v}}_i^T$:

$$\tilde{\mathbf{v}}_i^T = \arg \min_{\tilde{\mathbf{v}}_i^T} \left\| (\mathbf{r}^T \mathbf{F}_{\text{DFT}}) \circ (\tilde{\mathbf{v}}_i^T \mathbf{V}_1 \mathbf{F}_{\text{DFT}} - \boldsymbol{\delta}^T \mathbf{F}_{\text{DFT}}) \right\|_2, \quad (43)$$

where \mathbf{F}_{DFT} is the 2D DFT matrix, and $\boldsymbol{\delta}$ is a vector of the discrete delta function centered at the pixel corresponding to $\tilde{\mathbf{v}}_i^T$. Equations (42) and (43) are all the equations required to reconstruct the image, and replace Eq. (5).

Using the electric field model described in Section 2, we simulate imaging of an object in MATLAB. We compare the images obtained with the original and new algorithms in Fig. 3.

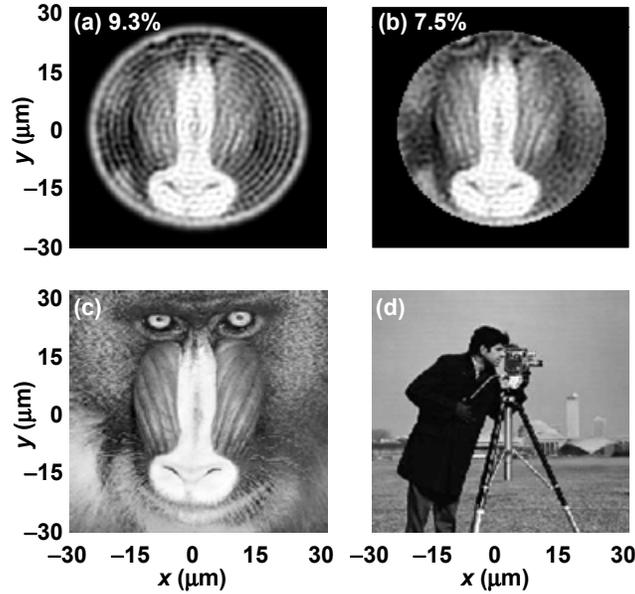


Fig. 3. Images reconstructed using OBR: (a) original algorithm and (b) new algorithm using an estimate of the object spatial frequency spectrum. Also shown are: (c) the object reflectance and (d) the object providing the estimated spatial frequency spectrum, which is unrelated to the actual object. In (a) and (b), values of the image error RMSD $\sigma(\tilde{w}_i)$ within the MMF core are given as a percent of the maximum image intensity. Intensity patterns and object are at 5 μm from the MMF fiber tip. The MMF has 50- μm core diameter and 0.45 NA. The number of sampling patterns is $16N$ (corresponding to $2 \times$ oversampling in the x - and y -dimensions for spots). Intensity patterns are sampled on a 256×256 grid.

We see that the image quality is substantially improved when using the new algorithm, which reduces the image error RMSD from 9.3% to 7.5% of the maximum image brightness – or equivalently, an increase in the peak signal-to-noise ratio from 20.6 to 22.5 dB. We do note that to use Eq. (43), it is necessary to know \mathbf{r} . This seems to present a paradox: if we know \mathbf{r} , we already have a perfect image of the object and imaging is unnecessary. In practice, however, the improvement shown in Fig. 3 can be obtained even if we have only an estimate of the spatial frequency spectrum of \mathbf{r} . The \mathbf{r} used in Eq. (43) is shown in Fig. 3(d), and differs substantially from the actual \mathbf{r} shown in Fig. 3(c).

We also briefly mention the issue of computation time. Equation (43) is a quadratic optimization problem with N variables, which can be solved efficiently by convex optimization methods. In practice, the amount of time to solve Eq. (43) on a personal computer is under 1 s. However, the number of separate $\tilde{\mathbf{v}}_i^T$ that must be solved is equal to L , the number of pixels in the final image (e.g., $L = 256^2$ for a 256×256 pixel image). Hence, real-time computation of Eq. (43) for $\tilde{\mathbf{V}}$ requires solving for the different $\tilde{\mathbf{v}}_i^T$ in parallel. We finally note that $\tilde{\mathbf{V}}$ is not unitary, and thus the noise amplification no longer depends on the singular values of \mathbf{UDV}^T but instead on the singular values of the matrix $\mathbf{UD}\tilde{\mathbf{V}}^T$, which are not the diagonal elements of \mathbf{D} . In practice, the changes to the noise amplification are minor.

4.3 Noise-reduced image reconstruction

Noise amplification can be further mitigated during image reconstruction by adding a constraint to the elements of each $\tilde{\mathbf{v}}_i^T$:

$$\tilde{\mathbf{v}}_i^T < \alpha \mathbf{d}^T \quad \forall i, \quad (44)$$

where \mathbf{d} is a row vector of the diagonal elements of \mathbf{D} and α is a real constant to be determined. Together with Eq. (43), this becomes:

$$\tilde{\mathbf{v}}_i^T = \arg \min_{\tilde{\mathbf{v}}_i^T} \left\| \left(\mathbf{r}^T \mathbf{F}_{\text{DFT}} \right) \circ \left(\tilde{\mathbf{v}}_i^T \mathbf{V}_1 \mathbf{F}_{\text{DFT}} - \delta^T \mathbf{F}_{\text{DFT}} \right) \right\|_2 \quad \text{s.t.} \quad \tilde{\mathbf{v}}_i^T < \alpha \mathbf{d}^T. \quad (45)$$

From Section 3, the noise amplification depends on the singular values of $\mathbf{U} \mathbf{D} \tilde{\mathbf{V}}^T$. When Eq. (51) is used instead of Eq. (43), the singular values are more equalized and thus the contribution to noise from low singular values is limited. However, by limiting the contribution to the image by intensity modes with low singular values, the resulting image has reduced resolution. The choice of α determines the tradeoff between noise amplification and resolution; when α is sufficiently high, Eq. (45) reduces to Eq. (43). We denote this value of α as α_{max} , given by $\alpha_{\text{max}} = \max(\tilde{\mathbf{V}} \tilde{\mathbf{D}}^{-1})$, where $\tilde{\mathbf{V}}$ is the solution to Eq. (43).

Using the electric field model described in Section 2, we simulate imaging of an object in MATLAB. We compare the images obtained using Eqs. (42) and (43) and Eqs. (44) and (45) in Fig. 4.

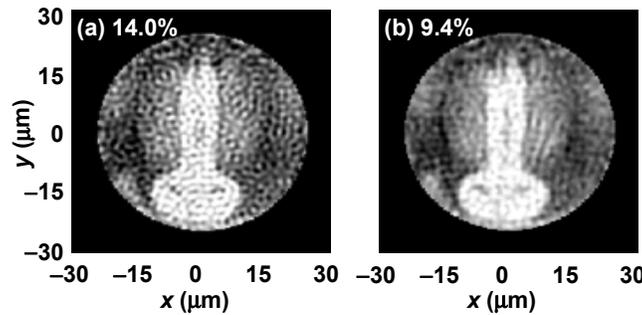


Fig. 4. Images reconstructed using OBR: (a) using original method of Eqs. (42) and (43), and (b) using reduced-noise method of Eqs. (44) and (45). The ratio of the noise constraint to the maximum noise constraint is $\alpha / \alpha_{\text{max}} = -10$ dB. The shot noise-limited SNR is

$\sigma_{\text{noise}} / \sqrt{\bar{\mathbf{p}}} = -22$ dB, where $\bar{\mathbf{p}}$ is the average power. Values of the image error RMSD $\sigma(\tilde{\mathbf{w}}_i)$ within the MMF core are given as a percent of the maximum image intensity. The intensity patterns and object are at $5 \mu\text{m}$ from MMF fiber tip. The MMF has $50\text{-}\mu\text{m}$ core diameter and 0.45 NA. The number of sampling patterns is $16N$ (corresponding to $2 \times$ oversampling in the x - and y -dimensions for spots). The intensity patterns are sampled on a 256×256 grid.

We observe that the image reconstructed using Eqs. (44) and (45) is less noisy, with a reduction in image error RMSD from 14.0% to 9.4%. For every noise SNR there is an optimal value of α that balances the image errors due to reduced resolution and noise, and thus minimizes the total image error. When \mathbf{r} is unknown, the optimal value of α is also unknown, but α can be chosen such that the image error RMSD is subjectively minimized, which is how α has been chosen here. Using Eqs. (44) and (45) reduces thermal, shot and intensity noise equally, and also reduces the error introduced due to the approximation of Eq. (6). Since this is always present even in the noiseless case, conservative (large) values of α lead to reductions in image error RMSD even when there is no noise.

5. Comparison of image quality

To directly compare the image error between images reconstructed using OBR and localized reconstruction, we simulate imaging an object using both methods in the presence of noise and plot the image error versus the power noise in Fig. 5. The apparatus and noise model are as described in Sections 2-3. The imaged object is shown previously in Fig. 3(d). In the plots of Fig. 5, we have normalized the noise values: at any SNR, $\sigma_{\text{intensity}} = \sigma_{\text{shot}} / \sqrt{\bar{p}} = \sigma_{\text{thermal}} / \bar{p}$, where \bar{p} denotes the average measured power for imaging using OBR. Equal values of σ_{thermal} , σ_{shot} , $\sigma_{\text{intensity}}$ and equal numbers of sampling patterns were used for both OBR and localized reconstruction. Results for both the maximum irradiance and total power constraints are shown in Fig. 5. For the maximum irradiance constraint the sampling patterns for OBR are random, while for the total power constraint the sampling patterns for OBR are spots, for reasons described in Section 4.1.

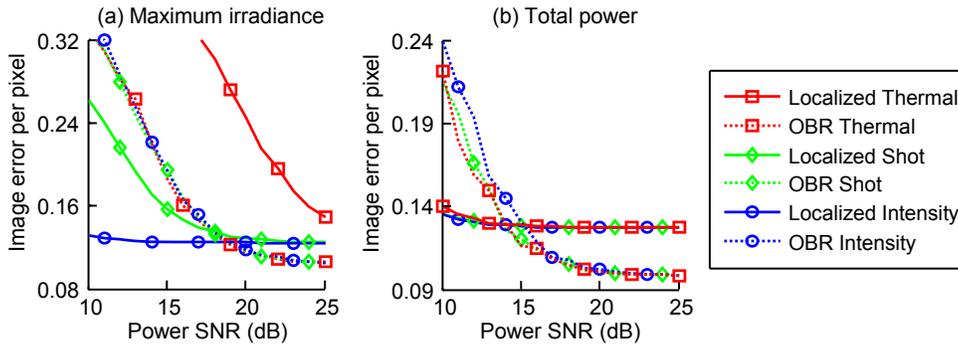


Fig. 5. Errors of images reconstructed using OBR (dashed lines) and localized reconstruction (solid lines) as a function of power noise levels under constraints on (a) maximum irradiance and (b) total power. The intensity patterns and object are at $5 \mu\text{m}$ from MMF fiber tip. The MMF has $50\text{-}\mu\text{m}$ core diameter and 0.45 NA. The number of sampling patterns is $16N$ (corresponding to $2 \times$ oversampling in the x - and y -dimensions for spots). Intensity patterns are sampled on a 256×256 grid.

The noise amplification curves shown in Fig. 5 behave as predicted based on Fig. 2. For low levels of noise, images reconstructed using OBR have a lower image error RMSD than those using localized reconstruction, because the former have higher resolution. For high levels of noise, images reconstructed using OBR have higher image error RMSD than localized reconstruction, due to noise amplification. Under the maximum irradiance constraint, the image error using OBR is less than that of localized reconstruction for all thermal noise levels, and for noises up to -18 dB and -19 dB for shot and intensity noises, respectively. When imaging using spot patterns under the total power constraint, the image error using OBR is less than that of localized reconstruction for noises up to -15 dB, -15 dB and -16 dB for thermal, shot and intensity noises, respectively.

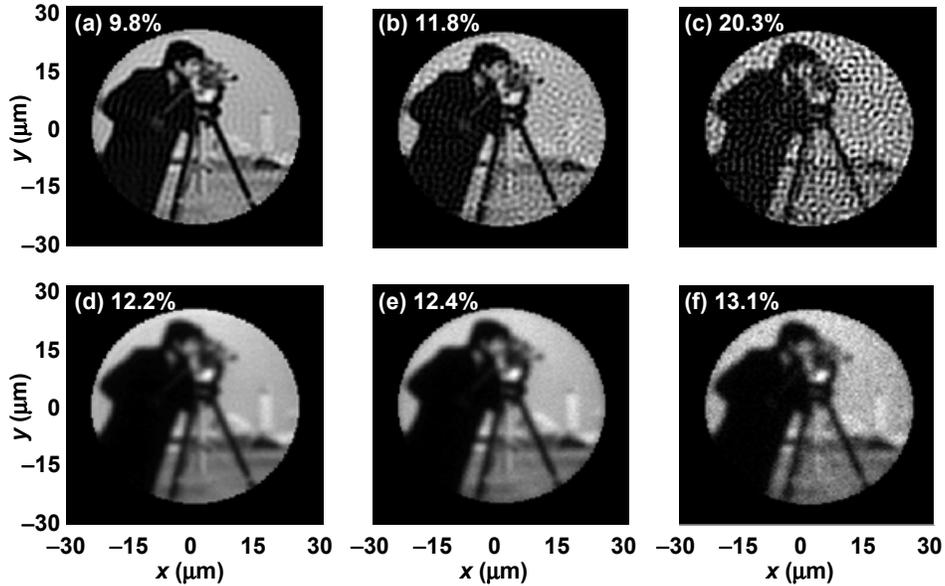


Fig. 6. Images reconstructed with OBR (a-c) and localized reconstruction (d-f) for various levels of shot noise. $\sigma_{\text{shot}} / \sqrt{P}$ is -25 dB for Figs. (a) and (d); -15 dB for Figs. (b) and (e); and -10 dB for Figs. (c) and (f). The sampling patterns are constrained by total power. Values of the image error RMSD $\sigma(\tilde{w}_i)$ within the MMF core are given as a percent of the maximum image intensity. The intensity patterns and object are at $5 \mu\text{m}$ from MMF fiber tip. The MMF has $50\text{-}\mu\text{m}$ core diameter and 0.45 NA. The number of sampling patterns is $16N$ (corresponding to $2 \times$ oversampling in the x- and y-dimensions for spots). The intensity patterns are sampled on a 256×256 grid.

In Fig. 6, we show images corresponding to three regimes on the plot of Fig. 5(b) as an example. These images are for the total power constraint and various levels of shot noise. OBR achieves superior resolution compared to localized reconstruction for low noise, as seen in Figs. 6(a) and 6(d), but noise amplification occurs for higher amounts of noise as seen in Figs. 6(c) and 6(f). In Fig. 6(a), for 25 dB SNR the image is still slightly noisy; this is due to the error introduced by the assumption made in Eq. (6).

6. Discussion

The advantages of OBR using random sampling patterns are faster calibration time and higher resolution compared to localized reconstruction, which come at the expense of noise amplification. By using the improved methods described in Section 4, an image can be reconstructed even with a moderate level of noise. If the noise level is high, averaging of the measured power values and reconstructed images may need to be done (as in [3]).

In [3], it was assumed that the apparatus can receive all the power reflected from the object. In Fig. 6, we demonstrate that this assumption introduces an error to the image. Future work may investigate exciting only the lower-order modes of high-NA MMFs for imaging, which is possible if the coupling between higher and lower order modes is limited [24]. By doing so, we would expect an increased fraction of the reflected power to couple back into the MMF due to the availability of higher-order modes, minimizing this error.

In this work, we have imaged objects that were placed $5 \mu\text{m}$ from the end of the MMF. As the separation between the MMF and the object increases, less power couples back into the MMF, and resolution decreases for both localized reconstruction and OBR, even in the

absence of noise. For a MMF with larger core radius (but smaller NA such that the total number of modes is preserved), this distance increases.

7. Conclusion

We have quantified the noise amplification for the optimization-based multi-mode fiber endoscopic imaging method originally described in [3], and described three improvements to reduce noise amplification. A comparison of images obtained using noise-reduced OBR and the original OBR yielded a 20% reduction in image error for the noiseless case, and a further 33% reduction in image error at 22 dB SNR for the shot noise-limited case. As a comparison, we simulated imaging an object using both OBR and localized reconstruction under two constraints: on the maximum irradiance or on the total power of the sampling patterns. When imaging using random sampling patterns under the maximum irradiance constraint, we found that the image error using OBR was less than that of localized reconstruction for all SNRs when thermal noise-limited, and down to 18 dB and 19 dB SNR for shot and intensity noise-limited cases, respectively. When using spot sampling patterns under the total power constraint, we found that the image error using OBR was less than that of localized reconstruction for SNRs down to 15 dB, 15 dB and 16 dB for thermal, shot and intensity noise-limited cases, respectively.

Acknowledgments

Support from the Samsung GRO Program is gratefully acknowledged.