

Computational endoscopy – a framework for improving spatial resolution in fiber bundle imaging

JOHN P. DUMAS¹, MUHAMMAD A. LODHI², BATOUL A. TAKI², WAHEED U. BAJWA², MARK C. PIERCE^{1,*}

¹Rutgers, The State Univ. of New Jersey, Dept. of Biomedical Engineering, 599 Taylor Road, Piscataway, NJ, USA 08854

²Rutgers, The State Univ. of New Jersey, Dept. of Electrical and Computer Engineering, 94 Brett Road, Piscataway, NJ, USA 08854

* Corresponding author: mark.pierce@rutgers.edu

Received XX Month XXXX; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

This letter presents a framework for computational imaging (CI) in fiber-bundle-based endoscopy systems. Multiple observations are acquired of objects spatially modulated with different random binary masks. Sparse-recovery algorithms then reconstruct images with more resolved pixels than individual fibers in the bundle. Object details lying within the diameter of single fibers are resolved, allowing images with 41,663 resolvable points to be generated through a bundle with 2,420 fibers. Computational fiber bundle imaging of micro- and macro-scale objects is demonstrated using fluorescent standards and biological tissues, including *in vivo* imaging of a human fingertip. In each case, CI recovers detail that conventional endoscopy does not provide.

<http://dx.doi.org/10.1364/OL.99.099999>

Endoscopic imaging provides real-time visualization of tissue at sites within the body. Larger instruments such as colonoscopes and gastroscopes (5-10 mm diameter) use digital CCD or CMOS sensors located at the distal tip. Smaller devices (below around 3 mm diameter) including bronchoscopes and pediatric scopes use coherent fiber optic bundles to transmit images to an eyepiece or sensor at the proximal end. Fiber bundles are also used for confocal, multi-photon, and epi-fluorescence endomicroscopy. Configurations that use a single optical fiber typically require a mechanism for distal beam scanning. Fiber bundles avoid this challenge, but the number of resolvable points is limited to the number of fibers within the bundle.

Post-processing techniques including Fourier filtering, spatial averaging, or interpolation can be applied to images acquired through a fiber bundle [1]. These methods can improve the visual appearance of images by removing the fiber bundle mask pattern, but the information content is still limited by the fiber packing density. Other techniques include image-compounding, combining multiple images as the distal end of the fiber bundle is laterally displaced between frames [2]. Bedard *et al.* use a similar concept,

but avoid the need to physically move the bundle by using a prism at the distal tip to spectrally encode spatially offset images that are captured in a snapshot multispectral image [3]. Image registration protocols have been developed to reconstruct a high-resolution image from a sequence of shifted, low-resolution frames for both known [4] and unknown [5, 6] displacements. Ravi *et al.* showed that deep learning algorithms further improve this approach [7]. Recently, Vyas *et al.* reported bundle-shifting confocal endomicroscopy with a nearly two-fold resolution improvement [8]. All the approaches discussed above improve the appearance of images acquired through fiber optic bundles, but each is based on individual fibers resolving only a single pixel of information.

Alternative approaches aiming to resolve multiple points include wavefront shaping for imaging through a single multimode fiber [9]. Ohayon *et al.* implemented wavefront shaping endoscopy for deep brain imaging, but fiber bending remains a challenge of this technique [10]. However, a recent exploration of step-index and graded-index multimode fibers suggests a path forward [9], while Shin *et al.* demonstrated a compressed sensing approach for confocal endomicroscopy that is insensitive to bending of the fiber bundle [11]. Unfortunately, the bundle is used only to deliver structured light to the sample, instead of collecting emitted light, making clinical implementation difficult at present.

This letter proposes a framework for resolution improvement using computational imaging (CI). Our approach requires collecting multiple images, or “observations”, of an object through different coded masks. We then use sparse-recovery algorithms to reconstruct an image that resolves multiple pixels *within* the diameter of each fiber in the bundle, producing images with a resolution determined by the density of elements in the mask rather than the number of fibers in the bundle.

We use a CI framework based on coded masks located at a conjugate image plane, which we previously evaluated when imaging without a fiber bundle [12]. We extended the mathematical forward model from this earlier framework to account for imaging through a fiber bundle. Since the inter-fiber gaps are opaque, information about regions of the object located behind these gaps is lost and the algorithms can only recover

details within the intra-fiber regions where light is collected. A binary $\{0, 1\}$ matrix, \mathbf{B} , is included within our system model to designate where spatial information about the sample is acquired. To construct this matrix, circle detection is used to set a value of 1 for pixels found to lie within fiber boundaries, and 0 for pixels in the inter-fiber regions. We choose the dimensions of our target reconstructed image, \mathbf{X} , to match those of an $m \times m$ mask, \mathbf{M}_k , placed at a conjugate image plane. The object is then imaged through k different random binary masks. Our forward model for a single $n \times n$ pixel camera observation \mathbf{Y}_k is defined as:

$$\mathbf{Y}_k = \mathbf{B} \odot \{\mathbf{D}_r [\mathbf{h} * (\mathbf{M}_k \odot \mathbf{X})] \mathbf{D}_c\} + \text{noise} \quad (1)$$

where $*$ denotes linear convolution, \odot denotes elementwise multiplication, \mathbf{h} is a Gaussian convolution kernel that accounts for system non-idealities [12], and \mathbf{D}_r and \mathbf{D}_c indicate two-dimensional downsampling. Equation (1) can be modified to fit the general format of $\mathbf{y}_k = \mathbf{A}_k \mathbf{x} + \text{noise}$ for sparse recovery, where a vector representation of the $m^2 \times 1$ object, \mathbf{x} , is solved for using multiple known system dependent measurement matrices, \mathbf{A}_k , and vectorized $n^2 \times 1$ observations, \mathbf{y}_k :

$$\mathbf{y}_k = \mathbf{T}_B \mathbf{D} \mathbf{T}_h \mathbf{T}_{M_k} \mathbf{x} + \text{noise} \quad (2)$$

where \mathbf{T}_{M_k} is an $m^2 \times m^2$ diagonal matrix with mask elements on the diagonal, \mathbf{T}_h is a block Toeplitz convolution matrix obtained from \mathbf{h} , \mathbf{D} combines downsampling in both directions, and \mathbf{T}_B is a diagonal matrix with the entries of \mathbf{B} on its diagonal. Thus, the system-dependent measurement matrix for each observation through our imaging setup can be expressed as $\mathbf{A}_k = \mathbf{T}_B \mathbf{D} \mathbf{T}_h \mathbf{T}_{M_k}$. To realize the zero-mean distribution characteristic of measurement matrices used in traditional sparse-recovery setups, we shift the mean value of our masks by subtracting 0.5 from all entries, and we shift the values in the associated observations by subtracting from each an image of a uniform object taken through a mask with all entries set to 0.5 [13].

The benchtop platform constructed to test this computational endoscopy approach (Fig. 1) is configured to image fluorescent objects. A 470 nm LED with a 429-474 nm bandpass filter (Semrock FF01-452/45-25) illuminates the sample. Sample fluorescence is imaged onto a digital micromirror device (DMD) embedded in a TI LightCrafter 6500 module that rapidly switches between sets of random binary masks, each containing 700×700 elements with 2×2 binning of DMD mirrors for an effective mask size of 350×350 . The modulated image of the object is then projected onto one end of a flexible fiber optic bundle (Schott

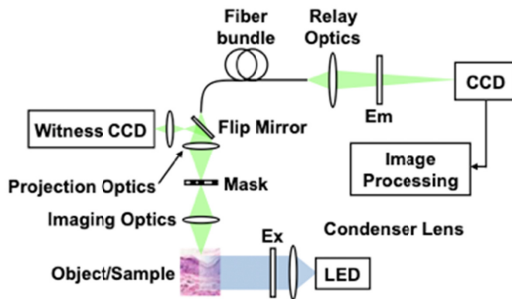


Fig. 1. System diagram for computational endoscopy. Synthetic objects are displayed directly at the mask plane.

Image Bundle 1621157) containing 18,000 individual multimode fibers, each 11.9 μm in diameter. Projection optics were selected to achieve a demagnification of $10\times$ from the DMD to the distal face of the bundle. This relates to 490,000 mask elements mapping onto a region of approximately 2,445 fibers. Relay optics image the proximal end face of the bundle through a 506-594 nm bandpass filter (Semrock FF01-550/88-25) and onto a CCD camera (Point Grey Research, GRAS-14S5M-C), such that approximately 100 pixels sample each fiber core.

Fig. 2 illustrates experimental results for a synthetic resolution target [Fig. 2(a)], generated in software, displayed with mask patterns imposed on the DMD, and then imaged through the fiber bundle via the projection and relay optics shown in Fig. 1. This allowed the system performance to be quantified by comparing pixel values obtained by computational reconstruction with the known pixel values of the synthetic object. The zoomed-in panels in Figs. 2(c)-(l) highlight details of the imaged target at the single fiber level. The inter-fiber cladding creates areas within the field-of-view (FOV) where object detail is blocked. Our CI approach does not aim to recover this missing data. Rather, we employ CI to recover intra-fiber details, producing images containing multiple resolved pixels within the diameter of individual fibers. The regions of the object that can be reconstructed are then illustrated in Figs. 2(e) and 2(f) as the original object with the bundle pattern overlaid. The raw image captured through the bundle is representative of conventional endoscopy. Figs. 2(g) and 2(h) show cropped regions of this image, illustrating that small object features are not resolved.

We reconstructed images by applying two separate sparse-recovery algorithms to the same set of observations. Algorithm choice is driven by the non-negativity constraint for reconstructed images. Previously, we used an algorithm based on Nesterov's proximal gradient (NPG) [14], which provided satisfactory results [12], but was computationally expensive due to a complex underlying implementation. Here, we use a computationally efficient least squares method [15] that imposes the non-negativity constraint and penalizes the L2-norm of the anisotropic total variation [16]. Images reconstructed with this total variation non-negativity-confined (TVNN) algorithm [Fig. 2(k) and 2(l)] are visually similar to reconstructions with the NPG algorithm [Fig. 2(i) and 2(j)], but are generated 8.6 times faster. For images processed in MATLAB v9.1 on a standard PC (Intel i7 processor @ 3.4 GHz, 16 GB RAM), the reconstruction process takes 110 s for TVNN and 945 s for NPG. Zoomed-in panels of the full 700×700 reconstructed images are presented in Fig. 2. Both algorithms recover intra-fiber details that are not captured with the conventional fiber bundle imaging method. Specifically, the blue highlighted area illustrates that corners and edges of line pairs can be resolved with CI and the red area illustrates that multiple line pairs are resolved within a single fiber. For each individual fiber, the CI reconstructions recover 17 pixels, a number dictated by the size of the displayed masks. While roughly 70 mask elements are mapped onto each fiber, 2×2 element binning results in approximately 17 unique elements per fiber. Fig. 2(b) shows line profiles across the single fibers indicated by the red arrows. Three maxima and two minima (corresponding to the two black lines) are resolved in the CI images. This generates approximately 17 resolved points within each fiber. In contrast, each fiber in the conventional bundle image provides only one point in the image.

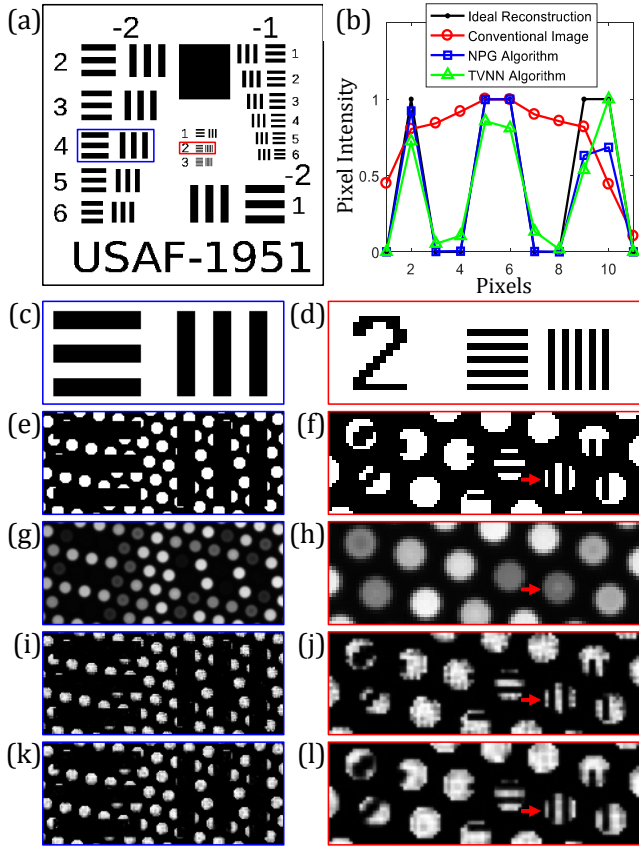


Fig. 2. Experimental results for the synthetic resolution target in (a). (b) Plot of line profiles across a single fiber (red arrows in (d-l)). (c, d) Zoom in of a 1.6×0.7 mm region (blue) and a 0.7×0.2 mm region (red) on the displayed target. (e, f) Ideal reconstruction of intra-fiber details. (g, h) Conventional bundle image. (i-l) CI reconstructions with 200 observations using (i, j) NPG algorithm and (k, l) TVNN algorithm.

Further quantitative measures demonstrate the relationship between the number of observations (masked images) and CI reconstruction accuracy [Fig. 3]. “Reconstruction error” is the normalized relative error obtained by comparing the displayed (true) image to the conventional and CI reconstructed images. Both algorithms outperform conventional imaging. Using more observations is beneficial for both NPG and TVNN reconstructions. However, beyond 80 observations, additional measurements do not significantly improve the reconstructed image quality.

Building on the quantitative evaluation in Figs. 2 and 3, we next imaged real objects with the platform illustrated in Fig. 1, including a mixture of $40 \mu\text{m}$ and $50 \mu\text{m}$ diameter fluorescent beads (Phosphorex Polystyrene Microspheres). A $10\times$ objective (Olympus Plan N) and 150 mm focal length achromatic lens were used to project an image of the beads onto the DMD. The “ground truth” image of the beads in Fig. 4(a) is a cropped region of a 980×1240 image taken with an independent “witness” camera positioned prior to the fiber bundle. When the same beads are imaged through the fiber bundle [Fig. 4(b)], the black inter-fiber cladding and pixilation due to individual fibers reduces overall brightness and image quality such that the edges of the beads are not well defined. Fig. 4(c) is the result of basic post processing of a single observation, applying bicubic interpolation to Fig. 4(b) to replace the fiber bundle mask pattern with values interpolated

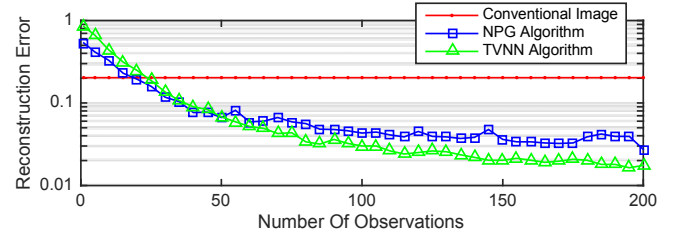


Fig. 3. Relative reconstruction error with increasing number of observations for NPG and TVNN on the same dataset shown in Fig. 2.

from the fibers. After interpolation the images appear more continuous and brighter, but integrity at the edges of the beads is not restored. Our CI method using the TVNN algorithm with 200 observations [Fig. 4(d)] captures the curved edges of the beads with better definition than the representative conventional fiber bundle image in Fig. 4(b). By removing the bundle mask pattern with interpolation of our CI reconstructed image, the outline of the beads becomes even clearer and bright spots within the interior can be distinguished. Comparing Fig. 4(c) and 4(e) clarifies that our CI method is closer to the ground truth image than the representative conventional fiber bundle image. Furthermore, it is easier to distinguish the $40 \mu\text{m}$ from the $50 \mu\text{m}$ beads with the CI method [Fig. 4(e)].

The middle row of Fig. 4 illustrates similar results, but for a different object on a different scale. Here, lens paper coated in fluorescent highlighter is imaged through the platform using a $4\times$ objective (Olympus Plan N) and a 150 mm focal length achromatic lens as the imaging optics. The red arrows indicate an area where individual paper fibers are better defined when CI is used with interpolation, compared to when interpolation is applied to the raw fiber bundle image. The blue arrows highlight where overlapping fibers can be distinguished in the CI image, but are not easily discerned in the conventional fiber bundle image.

The bottom row of Fig. 4 provides results for a trans-illuminated H&E stained histopathology section of normal colon tissue. The structure of the muscle tissue layers (top half of Fig. 4(k)), and crypts (bottom half of Fig. 4(k)) are characteristics of interest to pathologists. These features are visible in the ground truth image [Fig. 4(k)]. However, fine details are not discernable in the conventional fiber bundle image, with or without inter-fiber interpolation (Figs. 4(m) and 4(l), respectively). Computational imaging recovers tissue detail including crypt structure and organization [Fig. 4(o)]. The reduction in apparent brightness for CI reconstructed images (Figs. 4(j) and 4(n)) is partially due to the dark inter-fiber cladding regions and partially an artifact of system modeling errors where there is a mismatch between the mathematical model and physical system. Pixel values within individual fibers were similar to those in the ground truth image.

Many endoscopy applications work with a FOV on the order of millimeters to centimeters. Fig. 5 illustrates that computational endoscopy is scalable to macroscopic objects and potentially compatible with *in vivo* imaging. Fig. 5(a) shows a ground truth image of a human fingertip coated with fluorescent highlighter. At this scale, the computational endoscopy technique clearly recovers the narrow ridge patterns of the fingerprint [Fig. 5(c)] that are not resolved in conventional fiber bundle imaging [Fig. 5(b)].

In conclusion, we developed a framework to integrate CI with fiber bundle systems that is scalable to different FOVs and resolves

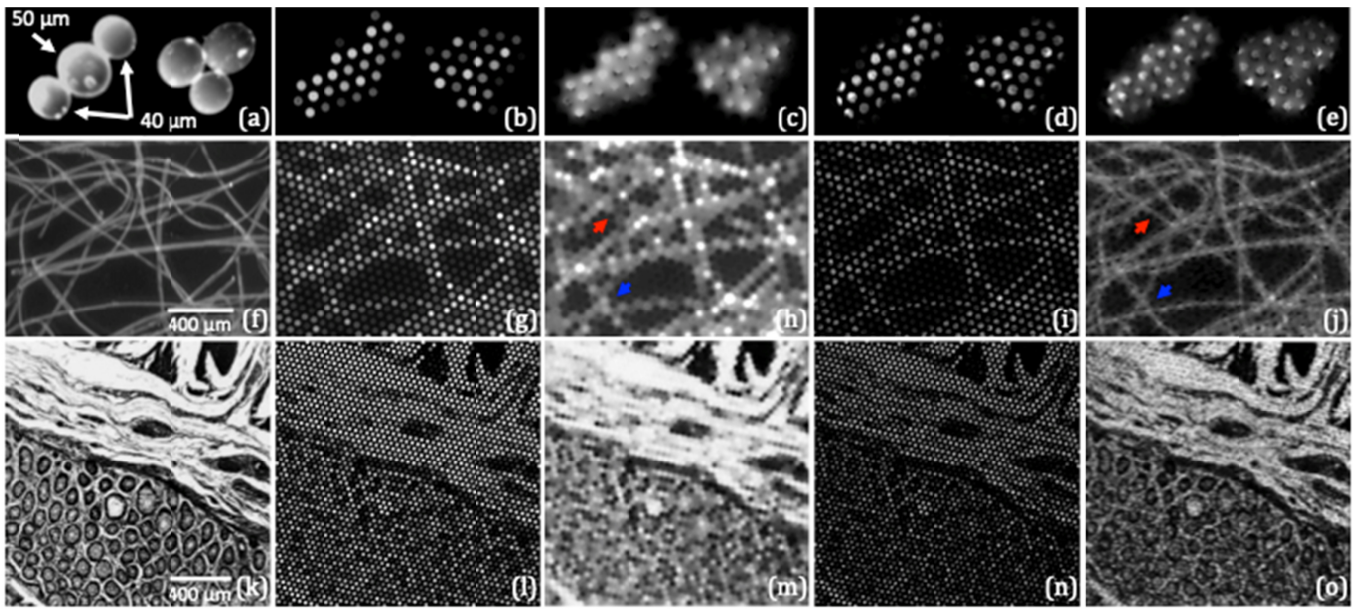


Fig. 4. (a, f, k) Ground truth images. (b, g, l) Single images captured through the bundle, representing conventional fiber-bundle imaging. (c, h, m) Bicubic interpolation on images (b, g, l). (d, i, n) CI reconstruction with 200 observations. (e, j, o) CI images (d, i, n) with bicubic interpolation.

17 pixels within the diameter of a single fiber. This approach does not attempt to recover detail from object regions within the inter-fiber cladding, but represents a step toward using sparse-recovery algorithms for resolution improvement in fiber-bundle-based endoscopes. However, two key challenges remain. First, image reconstruction time must be reduced, potentially by reducing the number of observations. At present, we do not use optimized algorithms to solve ill-conditioned CI problems. In our context this would mean recovering 17 points per fiber with fewer than 17 observations. Second, using a DMD for mask generation is impractical in the confined spaces inherent to endoscopy. One option for miniaturization would use a small glass window patterned with a chrome-etched mask that is rotated in front of the bundle tip with a miniature actuator. Using such a rotating mask pattern for CI in our platform when imaging the synthetic target in Fig. 2(a) produced an image with a reconstruction error of 0.0344 when using 100 observations, similar to results presented for random masks using the same number of observations (error of 0.0336). Furthermore, if the CI principles presented here are combined with spectral multiplexing, it may be possible to

reconstruct a high-resolution image in a single snapshot [17], eliminating the need for moving parts. We believe the introduction of CI concepts to endoscopy has the potential to advance progress toward *in vivo* optical biopsy.

Funding. National Science Foundation (NSF) (CCF-1453073, ECCS-1509260).

References

1. A. Shinde and M. V. Matham, J. Med. Imag. Heal. Inform. **4**, 203 (2014).
2. S. Park, M. Park, and M. Kang, in IEEE Signal Process. Mag. **20**, 21 (2003).
3. N. Bedard and T. S. Tkaczyk, J. Biomed. Opt. **17**, 080508 (2012).
4. M. Kyriash, R. Kester, R. Richards-Kortum, and T. Tkaczyk, Proc. SPIE **7558**, 07 (2010).
5. G. W. Cheon, J. Cha, and J. U. Kang, Opt. Lett. **39**, 4368 (2014).
6. J. Shao, W. Liao, R. Liang, and K. Barnard, Opt. Lett. **43**, 1906 (2018).
7. D. Ravi, A. B. Szczotka, D. I. Shakir, S. P. Pereira, and T. Vercauteren, Int. J. Comput. Assist. Radiol. Surg. **13**, 917 (2018).
8. K. Vyas, M. Hughes, B. G. Rosa, and G. Yang, Biomed. Opt. Express **9**, 4649 (2018).
9. O. Tzang, A. M. Caravaca-Aguirre, K. Wagner, and R. Piastun, Nat. Photonics **12**, 368 (2018).
10. S. Ohayon, A. M. Caravaca-Aguirre, R. Piastun, and J. J. DiCarlo, Biomed. Opt. Express **9**, 1492 (2018).
11. J. Shin, B. T. Bosworth, and M. A. Foster, Opt. Lett. **42**, 109 (2017).
12. J. P. Dumas, M. A. Lodhi, W. U. Bajwa, and M. C. Pierce, Opt. Express **25**, 29472 (2017).
13. D. Takhar, J. Laska, M. Wakin, M. Duarte, D. Baron, S. Sarvotham, K. Kelly, and R. Baraniuk, In SPIE Computational Imaging IV **6065**, 606509 (2006).
14. R. Gu, and A. Dogandzic in Proc. Asilomar Conf. Signals, Syst. Comput. **1662** (2014).
15. D. Needell, and R. Ward, SIAM J. Imag. Sci. **6**, 1035 (2013).
16. E. Candes and J. Romberg, 2005 [Online]. Available: <http://statweb.stanford.edu/~candes/l1magic/downloads/l1magic.pdf>
17. G. Arce, D. Brady, L. Carin, H. Arguello, and D. Kittle, IEEE Signal Process. Mag. **31**, 105 (2014).

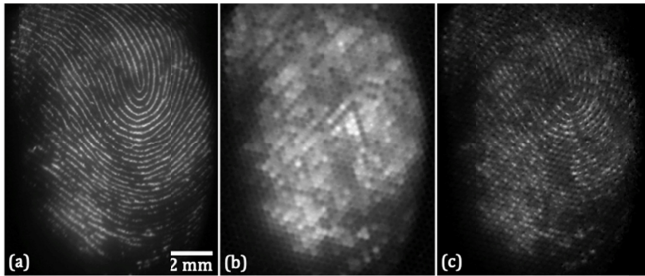


Fig. 5. Fluorescent highlighter applied to the fingertip. (a) Ground truth image. (b) Single image captured through the bundle with interpolation. (c) CI reconstruction with interpolation.

Full References

1. A. Shinde and M. V. Matham, "Pixelate removal in an image fiber probe endoscope incorporating comb structure removal methods," *J. Med. Imag. Health Inform.* 4, 203–211 (2014).
2. S. C. Park, M. K. Park, M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.* 20(3), 21–36 (2003).
3. N. Bedard and T. S. Tkaczyk, "Snapshot spectrally encoded fluorescence imaging through a fiber bundle," *J. Biomed. Opt.* 17(8), 080508 (2012).
4. M. Kyrish, R. Kester, R. Richards-Kortum, and T. Tkaczyk, "Improving spatial resolution of a fiber bundle optical biopsy," *Proc. SPIE 7558, Endoscopic Microscopy V*, 755807, 755807-9 (2010).
5. G. W. Cheon, J. Cha, and J. U. Kang, "Random transverse motion-induced spatial compounding for fiber bundle imaging," *Opt. Lett.* 39, 4368–4371 (2014).
6. J. Shao, W. C. Liao, R. Liang, and K. Barnard, "Resolution enhancement for fiber bundle imaging using maximum a posteriori estimation," *Opt. Lett.* 43, 1906–1909 (2018).
7. D. Ravi, A. B. Szczotka, D. I. Shakir, S. P. Pereira, and T. Vercauteren, "Effective deep learning training for single-image super-resolution in endomicroscopy exploiting video-registration-based reconstruction," *International journal of computer assisted radiology and surgery*, 1-8 (2018).
8. K. Vyas, M. Hughes, B. G. Rosa, and G. Z. Yang, "Fiber bundle shifting endomicroscopy for high-resolution imaging," *Biomed. Opt. Express*, 9(10), 4649–4664 (2018).
9. O. Tzang, A. M. Caravaca-Aguirre, K. Wagner, and R. Piestun, "Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres," *Nat. Photonics* 12, 368–374 (2018).
10. S. Ohayon, A. Caravaca-Aguirre, R. Piestun, and J. J. DiCarlo, "Minimally invasive multimode optical fiber microendoscope for deep brain fluorescence imaging," *Biomed. Opt. Express* 9, 1492–1509 (2018).
11. J. Shin, B. T. Bosworth, and M. A. Foster, "Compressive fluorescence imaging using a multi-core fiber and spatially dependent scattering," *Opt. Lett.* 42(1), 109–112 (2017).
12. J. P. Dumas, M. A. Lodhi, W. U. Bajwa, and M. C. Pierce, "From modeling to hardware: an experimental evaluation of image plane and Fourier plane coded compressive optical imaging," *Opt. Express* 25, 29472–29491 (2017).
13. D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," in *Computational Imaging IV*, 6065, (San Jose, CA, USA), p. 606509, SPIE, 2006.
14. R. Gu and A. Dogandžić, "A fast proximal gradient algorithm for reconstructing nonnegative signals with sparse transform coefficients," in *Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers (IEEE, 2014)*, pp. 1662–1667.
15. D. Needell and R. Ward, "Stable image reconstruction using total variation minimization," *SIAM J. Imaging Sci.* 6, 1035–1058 (2013).
16. E. J. Candès and J. Romberg. *ℓ1-MAGIC: Recovery of Sparse Signals via Convex Programming*, 2005. Available at <http://statweb.stanford.edu/~candes/l1magic/downloads/l1magic.pdf>
17. G. Arce, D. Brady, L. Carin, H. Arguello, and D. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Process. Mag.* 31(1), 105–115 (2014).