

Evaluation of computational endomicroscopy architectures for minimally-invasive optical biopsy

John P. Dumas^a, Muhammad A. Lodhi^b, Waheed U. Bajwa^b, Mark C. Pierce^{*a}

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

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

ABSTRACT

We are investigating compressive sensing architectures for applications in endomicroscopy, where the narrow diameter probes required for tissue access can limit the achievable spatial resolution. We hypothesize that the compressive sensing framework can be used to overcome the fundamental pixel number limitation in fiber-bundle based endomicroscopy by reconstructing images with more resolvable points than fibers in the bundle. An experimental test platform was assembled to evaluate and compare two candidate architectures, based on introducing a coded amplitude mask at either a conjugate image or Fourier plane within the optical system. The benchtop platform consists of a common illumination and object path followed by separate imaging arms for each compressive architecture. The imaging arms contain a digital micromirror device (DMD) as a reprogrammable mask, with a CCD camera for image acquisition. One arm has the DMD positioned at a conjugate image plane (“IP arm”), while the other arm has the DMD positioned at a Fourier plane (“FP arm”). Lenses were selected and positioned within each arm to achieve an element-to-pixel ratio of 16 (230,400 mask elements mapped onto 14,400 camera pixels). We discuss our mathematical model for each system arm and outline the importance of accounting for system non-idealities. Reconstruction of a 1951 USAF resolution target using optimization-based compressive sensing algorithms produced images with higher spatial resolution than bicubic interpolation for both system arms when system non-idealities are included in the model. Furthermore, images generated with image plane coding appear to exhibit higher spatial resolution, but more noise, than images acquired through Fourier plane coding.

Keywords: Compressive sensing, computational imaging, endomicroscopy.

1. INTRODUCTION

Medical endoscopes commonly use coherent fiber-optic bundles to facilitate *in vivo* imaging. Fiber bundles can be compact and highly flexible, allowing for excellent tissue access. However, these components also suffer from limited spatial resolution due to the finite number of discrete fibers which serve as individual pixels within the bundle. While small, densely packed fibers are required for achieving a high number of resolvable points per unit area, this is technically challenging to fabricate. Closely spaced fibers also suffer from increased fiber-to-fiber coupling, or cross-talk, particularly at near-infrared wavelengths¹.

A new approach is needed in order to overcome the resolution limitations of fiber bundle based imaging without increasing the outer diameter of the probe. To address this need, we are integrating signal processing concepts from the compressive sensing (CS) field with traditional endomicroscopy systems. We first consider the bundle as effectively a low-resolution sensor array in which the “ n ” individual fibers act as single pixels. Our approach for “*computational endomicroscopy*” then involves the addition of a coded mask into the imaging path with “ N ” spatial elements, where $N > n$ (Fig. 1) and we seek to recover an image with N pixels instead of n . For our initial work reported in this paper, we simplify the endomicroscopy system by replacing the fiber bundle with a CCD array. Future studies will consider the additional challenges introduced by true fiber bundle imaging such as the disordered organization of fibers within the bundle and mapping of mask elements to multiple fibers.

*mark.pierce@rutgers.edu; phone 1 848 445-6570; fax 1 732 445-3753

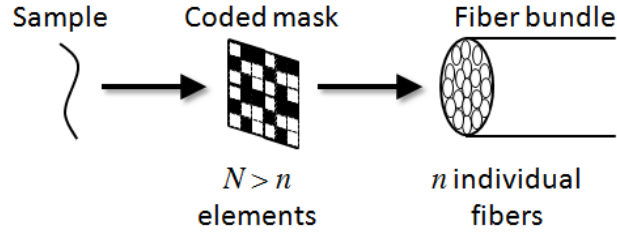


Figure 1. Computational endomicroscopy based on CS principles uses a coded mask with N programmable elements to modulate the scene before imaging through a fiber bundle containing n individual fibers. Following acquisition of multiple camera images, each with a different (known) mask pattern, application of an inverse system model aims to recover an image of the sample with N resolvable points.

The CS framework can be implemented in applications where the desired signal is sparse in some known transform domain². When integrated into an optical imaging system, where images are often sparse in wavelet or other domains, CS techniques make it possible to generate images with more resolvable points than are physically present in the optical sensor. An extreme example of CS-based imaging is the “single-pixel camera”, which uses a single photodetector to reconstruct images with several hundred thousand pixels³. There is, of course, a trade-off with temporal resolution as many sequential measurements of the object are required, each measuring a differently modulated projection of the object. In single-pixel camera architectures, modulation is typically achieved using a spatial light modulator (SLM) configured to impart a high-resolution intensity or phase modulation onto the object field, with the resulting light signal integrated at a single photodetector. Extensions of the single-pixel camera integrate focal plane arrays to address the limitations of imaging speed and light collection efficiency inherent in this highly-sequential measurement method by implementing similar CS principles in a parallel manner⁴⁻⁷.

Several theoretical mathematical models for CS have been proposed, some of which have been implemented physically in optical imaging systems. Generally, these systems differ based on the nature of the modulation (amplitude or phase) and the location of the modulating element within the optical path. Most often, a SLM is placed at a conjugate-image plane between the object and the sensor to modulate the object’s intensity as a function of position (Figure 2(a)). Other architectures place the SLM at a conjugate-Fourier plane (Figure 2(b)). This latter configuration opens the possibility of modulating both the phase and/or magnitude of the spatial frequencies contained within a scene. CS is also possible in variations of these architectures where the SLM is not placed directly at an image or Fourier plane. More generally, CS systems that use SLMs as coded apertures in the imaging path induce a modified point spread function in the system that is exploited during the CS reconstruction process. One of the most successful applications of the coded aperture architecture is in hyperspectral imaging with the CASSI project and its variations⁸⁻¹⁰. Alternatively, the SLM can be placed in the illumination path, either at an image plane for patterned illumination⁵ or at a Fourier plane as is done in Fourier ptychography systems^{11,12}.

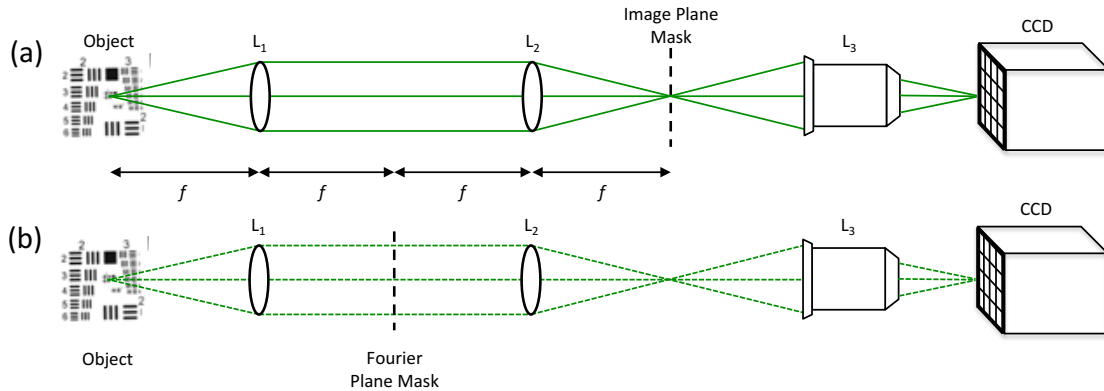


Figure 2. Schematic diagram of two CS architectures illustrating mask placement for (a) image plane coding and (b) Fourier plane coding.

Keeping our end goal of translating CS architectures to applications in endomicroscopy in mind, we are limited in which CS architectures we can pursue in practice. To maintain the possibility of fluorescence imaging, architectures that require phase-modulated illumination, such as traditional Fourier ptychography, are not feasible, because fluorophores are not responsive to phase variation of the excitation waves¹¹. Furthermore, since the light emitted from fluorescent samples is spatially incoherent, architectures that use phase modulation in the imaging path are not feasible as they are more appropriate for coherent imaging¹³. Architectures that encode the object at a conjugate image plane, whether in the illumination or imaging path, are not as dependent on the spatial coherence of the light setting because they only modulate an object's magnitude, not its phase profile. However, in an endomicroscopy system where trans-illumination is not plausible, a patterned illumination architecture becomes more complicated than using a transparent SLM in the imaging path. With the SLM in the illumination path, the angle of incidence must be considered and stricter requirements are placed on the optics, including restricted F-number and the degree of telecentricity at the object plane. Therefore, we investigated two distinct CS architectures with amplitude-modulating SLMs in the imaging path that are practical candidates for future implementation in endomicroscopy. We developed a benchtop platform to directly compare image-plane to Fourier-domain modulation with the goal of evaluating computational and hardware challenges in each architecture.

2. METHODS

2.1 Experimental platform

A benchtop platform was assembled consisting of two arms for separate image- and Fourier-plane-coding (Figure 3). The platform uses two identical digital micromirror devices (DMDs) (Texas Instruments, DLP® LightCrafter™ 6500 Evaluation Module) to display two-dimensional grayscale mask patterns. The DMDs are positioned at a conjugate image plane (IP), and a conjugate Fourier plane (FP) in the two arms of the system. When imaging with the IP arm, a cold white LED is used to illuminate the object. However, the discrete, tilting mirrors of a DMD can cause these components to act as blazed diffraction gratings when placed at a conjugate Fourier plane¹⁴. This behavior results in a blurred image at the camera under broadband illumination as different wavelengths exit the DMD at different angles. To generate a spatially incoherent light setting for imaging with the FP arm, we instead illuminate objects with collimated light from a single-mode fiber-coupled HeNe laser ($\lambda = 633 \text{ nm}$), transmitted through a spinning ground glass diffuser (Thorlabs, DG10-600-MD) to remove the appearance of speckle. A flip mirror is used to direct the image of the object through either platform arm. Both arms consist of lenses aligned in a 4f system where the focal length for $L_1 = L_2 = L_4 = 150 \text{ mm}$ (Thorlabs, AC254-150-A). A 10x objective L_3 (Olympus, UPLFLN 10X2) images each DMD onto a 14-bit CCD array with 1384×1036 pixels, each $6.45 \mu\text{m}$ square (Point Grey Research, GRAS-14S5M-C). An image can also be directed to a high-resolution “witness” camera with 3856×2764 pixels, each $1.67 \mu\text{m}$ square (The Imaging Source, DMK 24UJ003) to capture a ground truth image of the object.

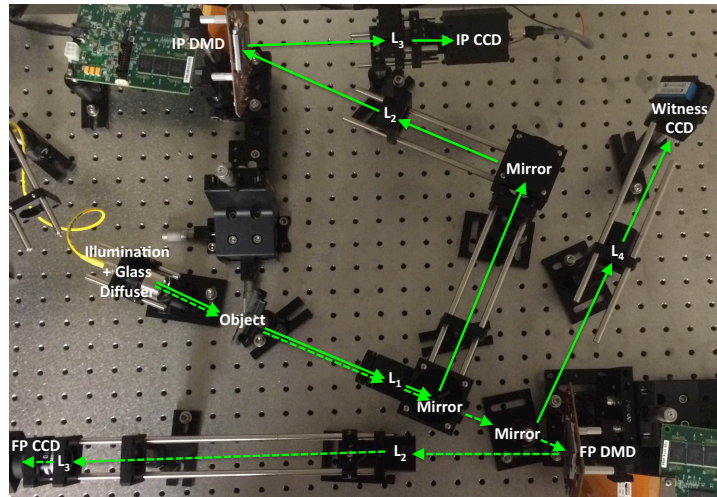


Figure 3. Photo of the benchtop computational imaging platform displaying beam paths for the image plane arm (solid line) and the Fourier plane arm (dotted line). The high-resolution “witness” camera takes ground truth images of the object for comparison to computationally reconstructed images.

2.2 Mathematical Models

The hardware implementation of each arm is a physical realization of two different mathematical models. The image-plane-coding arm first combines the object X with a mask M to form a pointwise modulated version ($M \odot X$) of the object intensity, where $\odot$ denotes elementwise multiplication of entries. This modulated object is then imaged onto the camera to obtain a single observation Y . In a perfectly aligned system, Y depends only on the modulated object and the element-to-pixel ratio d , which represents the number of mask elements that are mapped onto a single pixel. We achieve a element-to-pixel ratio of 16 by mapping 230,400 mask elements onto 14,400 camera pixels. The model for an ideal system can thus be represented as:

$$Y = D_n (M \odot X) \quad (1)$$

where D_n represents n -dimensional downsampling. The mask pattern M is changed several times and the corresponding observations (Y 's) are recorded, generating a linear system of equations with unknown values X . We then use optimization-based CS reconstruction methods to recover X from the measured Y 's and the known mask patterns (M 's), generating images with 230,400 pixels. For the IP arm, the reconstruction algorithm is based on Nesterov's proximal-gradient (NPG) method¹⁵, whereas for the FP arm, it is based on two-step iterative shrinkage/thresholding (TwIST) method¹⁶. Further discussion of different CS reconstruction methods is beyond the scope of this paper and will be reported in a later study.

Unfortunately, conditions are not ideal in practice due to alignment imperfections, optical aberrations, and hardware limitations. Axial or lateral misalignment can lead to mask elements being mapped onto multiple camera pixels (Figure 4(b)). Distortion in the optical path can create non-uniformities in mapping of mask elements to the camera (Figure 4(c)). Other optical aberrations can affect the point spread function, causing image blurring at the camera. Collectively, these effects result in light from individual mask elements being captured by camera pixels lying beyond the geometrically mapped pixel (Figure 4(f)).

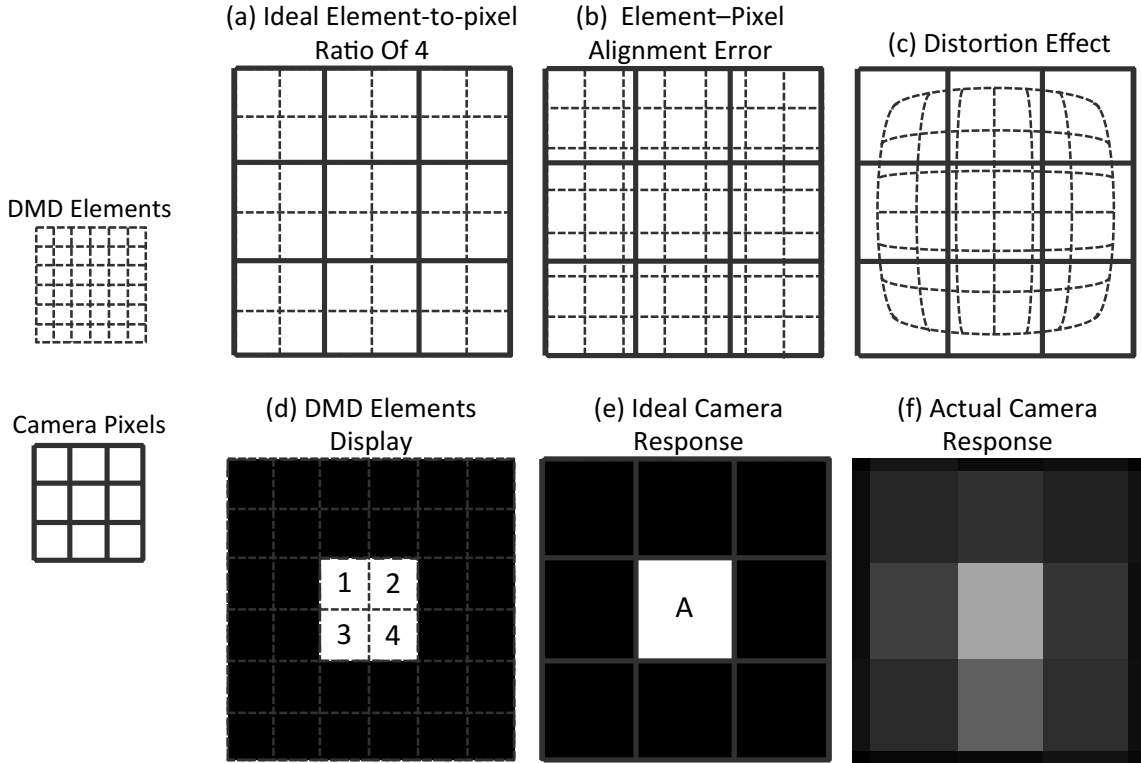


Figure 4. Illustration of non-ideal optical behavior in the IP arm. (a) Ideal 4× element-to-pixel ratio; exactly 4 DMD (mask) elements are imaged onto each pixel of the camera. (d, e) Ideally, all light from mask elements labeled 1–4 is then collected by camera pixel “A”. While element–pixel alignment errors (b) and distortion effects (c) can be minimized, the experimental camera response (f) still incurs some light leaking onto neighboring pixels.

To account for these non-idealities, we illuminate the system with a series of known object patterns (M) and directly measure the responses (Y) at each camera. This process allows us to estimate the effect of these non-idealities in a principled, quantitative manner. Once characterized, this effect is then integrated into our model using a single term H and equation (1) is thus updated to:

$$Y = D_n[H \otimes (M \odot X)] \quad (2)$$

where $\odot$ denotes convolution. Under ideal conditions, H is a $\sqrt{d} \times \sqrt{d}$ matrix with all entries equal to $1/d$ (where d is the element-to-pixel ratio), indicating an equal and ideal mapping of light from each mask element to a single camera pixel (Figure 4(d,e)). In practice, however, H should be selected as a larger square matrix to account for contributions from any number of mask elements to individual pixels. As H is alignment and system dependent, it must be experimentally determined following alignment adjustments and for different systems.

The preceding discussion also applies to the Fourier-plane-coding arm, with the major difference being the mask placement at a conjugate Fourier plane to modulate the spatial frequencies of the object. In hardware this is achieved by positioning the DMD at the Fourier plane in the 4f system where lens L1 implements a Fourier transform on the object. In this Fourier optics setting, the point spread function of the system is determined by the mask and can be expressed as

$|\bar{\mathfrak{F}}(M)|^2$, where $\bar{\mathfrak{F}}$ represents the inverse Fourier transform¹³. Therefore, our model for the Fourier-plane-coding arm can be expressed as a convolution of the magnitude-squared of the inverse Fourier transform of the mask with the object, giving:

$$Y = D_n \left[H \otimes \left(|\bar{\mathfrak{F}}(M)|^2 \otimes X \right) \right] \quad (3)$$

where H , in this case, is estimated as a Gaussian filter.

3. RESULTS

We imaged a printed USAF 1951 resolution target with both the witness camera and each CS arm of the system, focusing on a 120 x 120 pixel region from each low-resolution camera. Figure 5(a) shows the ground truth high-resolution image from the witness camera with an accompanying line profile across the Group 3 elements (along the red line) to highlight fine structure. Loss of fine details is observed when a single image is taken with the low-resolution cameras (Figure 5(b,c)). Post-acquisition image enhancement techniques, such as bicubic interpolation on the low-resolution image, are unable to recover much of the lost detail (Figure 5(d,g)). Using 50 sequential masked measurements, our CS reconstruction of the target scene with the ideal models shows a slight improvement of overall image quality spatial resolution for the IP arm, but not the FP arm (Figure 5(e,h)). The best results for each arm are produced when the model is corrected to account for system specific imperfections. Figure 5(f,i) shows sharper edges and more detail, particularly in the Group 3 elements, than bicubic interpolation. However, reconstructions using the modified model have a textured, grainy appearance, which is most likely an artifact of the reconstruction algorithm.

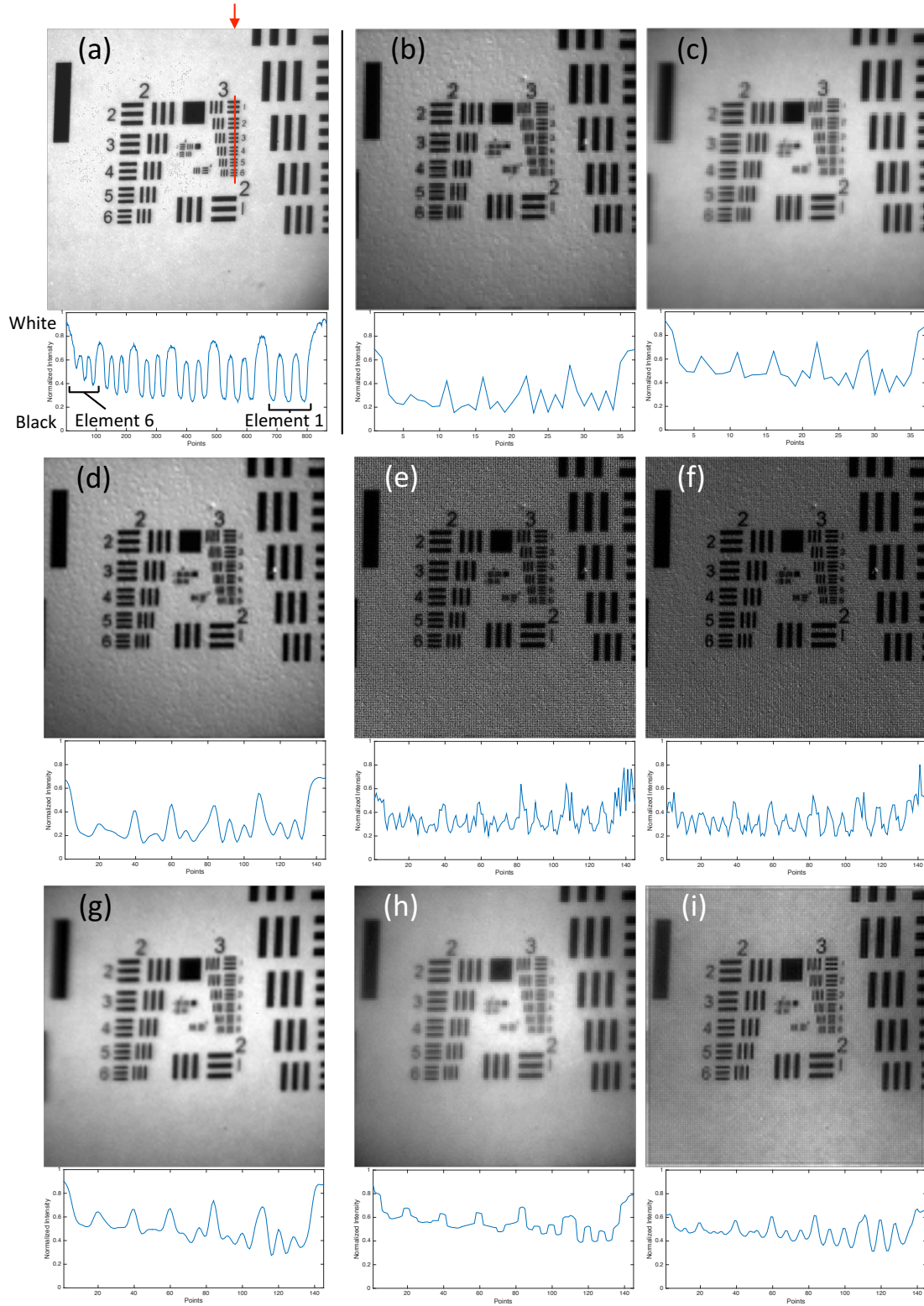


Figure 5. Experimental results showing line profiles and images from (a) the high-resolution “witness” camera, alongside the low-resolution cameras in the (b) IP arm, and (c) FP arm. Processed images from the image plane arm using (d) bicubic interpolation; computational imaging with (e) the ideal system model, and (f) the modified system model. Processed images from the Fourier plane arm using (g) bicubic interpolation, (h) the ideal system model, and (i) the modified system model.

4. CONCLUSIONS

Based on qualitative assessment of reconstructed images and line profiles, accounting for system specific non-idealities in CS models is a necessary step when reconstructing high-resolution images of a scene. We have shown that using magnitude-only coded masks at either a conjugate image or Fourier plane can be implemented to overcome the physical resolution limitations of a sensor. Using the CS framework, we recover more detail than applying bicubic interpolation to a single image, when a system-specific CS model is used. Furthermore, comparing Figures 5f and 5i, with the same number of observations our IP arm can recover more detail than our FP arm. However, reconstructions from the IP arm exhibit lower mean intensity and more pronounced artifacts, leading to generally noisier reconstructed images.

Future goals involve evaluating different optimization-based CS algorithms and modifying the mathematical models to better match the experimental system in order to reduce these reconstruction artifacts. We also aim to improve image quality and reduce the number of measurements required to generate a single image by experimenting with different mask patterns before translating our CS architecture to a fiber bundle based endomicroscopy platform.

ACKNOWLEDGEMENTS

This research was funded by the NSF (CCF-1453073, ECCS- 1509260), and ARO (W911NF-14-1-0295).

REFERENCES

- [1] Chen, X., Reichenbach, K. L., and Xu, C., "Experimental and theoretical analysis of core-to-core coupling on fiber bundle imaging," *Optics Express*, 16(26):21598-21607 (2008).
- [2] Baraniuk, R. G., "Compressive sensing," *IEEE Signal Processing Magazine* 24, 118-121 (2007).
- [3] Duarte, M. F., Davenport, M. A., Takhar, D., Laska, J. N., Sun, T., Kelly, K. E., and Baraniuk, R. G., "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, 25(2), 83 (2008).
- [4] Dumas, J. P., Lodhi, M. A., Bajwa, W. U., and Pierce, M. C., "Computational imaging with a highly parallel image-plane-coded architecture: challenges and solutions" *Optics Express*, 24(6), 6145-6155 (2015).
- [5] Shepard, R. H., Fernandez-Cull, C., Raskar, R., Shi, B., Barsi, C., and Zhao, H., "Optical design and characterization of an advanced computational imaging system," *Proc. SPIE* 9216; 92160A (2014).
- [6] Ke, J., and Lam, E. Y., "Object reconstruction in block-based compressive imaging," *Optics Express*, 20(20), 22102-22117 (2012).
- [7] Chen, H., Asif, M. S., Sankaranarayanan, A. C., and Veeraraghavan, A., "FPA-CS: Focal plane array-based compressive imaging in short-wave infrared," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2358-2366 (2015).
- [8] Arguello, H., and Arce, G. R., "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions on Image Processing*, 23(4), 1896-1908 (2014).
- [9] Galvis, L., Arguello, H., and Arce, G. R., "Coded aperture design in mismatched compressive spectral imaging," *Applied Optics*, 54(33), 9875-9882 (2015).
- [10] Rueda, H. F., Arguello, H., and Arce, G. R. "On super-resolved coded aperture spectral imaging," *Proc. SPIE* 8743 (2013).
- [11] Dong, S., Nanda, P., Shiradkar, R., Guo, K., and Zheng, G., "High-resolution fluorescence imaging via pattern-illuminated Fourier ptychography," *Optics Express*, 22(17), 20856-20870 (2014).
- [12] Tian, L., Li, X., Ramchandran, K., and Waller, L., "Multiplexed coded illumination for Fourier Ptychography with an LED array microscope," *Biomedical Optics Express*, 5(7), 2376-2389 (2014).
- [13] Willett, R. M., Marcia, R. F., and Nichols, J. M., "Compressed sensing for practical optical imaging systems: a tutorial," *Optical Engineering*, 50(7), 072601-072601 (2011).
- [14] Dudley, D., Duncan, W., and Slaughter, J., "Emerging digital micromirror device (DMD) applications," *Proc. SPIE* 4985, 14 (2003).
- [15] Gu, R., and Dogandzic, A., "A fast proximal gradient algorithm for reconstructing nonnegative signals with sparse transform coefficients," *Proc. Asilomar Conference on Signals, Systems, and Computers* 1662-1667 (2014).
- [16] Bioucas-Dias, J. M., and Figueiredo, M. A., "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image processing*, 16(12), 2992-3004 (2007).