

# Optical Engineering

[OpticalEngineering.SPIEDigitalLibrary.org](http://OpticalEngineering.SPIEDigitalLibrary.org)

## **Lensless coded-aperture imaging with separable Doubly-Toeplitz masks**

Michael J. DeWeert  
Brian P. Farm

# Lensless coded-aperture imaging with separable Doubly-Toeplitz masks

Michael J. DeWeert\* and Brian P. Farm

BAE Systems Spectral Solutions, 841 Bishop Street Suite 300, Honolulu, Hawaii 96813, United States

**Abstract.** In certain imaging applications, conventional lens technology is constrained by the lack of materials which can effectively focus the radiation within a reasonable weight and volume. One solution is to use coded apertures—opaque plates perforated with multiple pinhole-like openings. If the openings are arranged in an appropriate pattern, then the images can be decoded and a clear image computed. Recently, computational imaging and the search for a means of producing programmable software-defined optics have revived interest in coded apertures. The former state-of-the-art masks, modified uniformly redundant arrays (MURAs), are effective for compact objects against uniform backgrounds, but have substantial drawbacks for extended scenes: (1) MURAs present an inherently ill-posed inversion problem that is unmanageable for large images, and (2) they are susceptible to diffraction: a diffracted MURA is no longer a MURA. We present a new class of coded apertures, separable Doubly-Toeplitz masks, which are efficiently decodable even for very large images—orders of magnitude faster than MURAs, and which remain decodable when diffracted. We implemented the masks using programmable spatial-light-modulators. Imaging experiments confirmed the effectiveness of separable Doubly-Toeplitz masks—images collected in natural light of extended outdoor scenes are rendered clearly. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.54.2.023102](https://doi.org/10.1117/1.OE.54.2.023102)]

Keywords: coded apertures; computational imaging; lensless imaging; software-defined optics; spatial-light-modulators; Toeplitz.

Paper 141303P received Aug. 19, 2014; accepted for publication Dec. 30, 2014; published online Feb. 3, 2015.

## 1 Introduction

Coded-aperture imaging was introduced by Dicke<sup>1</sup> and Ables<sup>2</sup> in the 1960s, and the field developed rapidly for astronomical x-ray and gamma-ray imaging. Coded-aperture imagers extend the pinhole camera concept, which requires no lenses, has unlimited depth of focus, and can image radiation of any wavelength. By placing a large number of pinholes in a common aperture plane, the light-gathering capability is greatly increased. Improved sensitivity comes at the price of having the focal plane record a multiplex of overlapping images, requiring algorithmic reconstruction to render a clear image. Thus, coded-aperture imaging is a subset of the general field of computational imaging.

Some limitations of a random pinhole distribution, primarily numerical instability due to the existence of zeros in the inverse-mask Fourier-domain transfer function, were mitigated with the development of coded apertures with uniformly redundant arrays (URAs)<sup>3</sup> and modified URAs (MURAs)<sup>4</sup> of pinholes. By placing the pinholes in prime-number-based patterns determined by sampling theory, it was thought that the near-delta-function system point-spread functions (PSFs) could be achieved (after deconvolving the distribution with appropriate filter functions), greatly improving the image reconstruction.

Use of coded-aperture imaging for general terrestrial applications poses two challenges. First, strictly speaking, the information-theoretic advantages of URAs and MURAs apply only for radiation wavelengths short enough that diffraction is negligible, since a diffracted MURA is no longer a

MURA. Thus, coded-aperture imaging was applied mainly for x-ray and gamma-ray astronomies. Second, MURAs have limited effectiveness for extended scenes in general, regardless of radiation wavelength, due to the fundamentally ill-posed nature of reconstructing an extended scene from a finite-sized coded-aperture image. Thus, exploiting the potential of coded-aperture imaging requires addressing both diffraction and ill-posed reconstruction with very fast and efficient decoding methods. The work presented in this paper shows how to accomplish that goal.

### 1.1 Objective: Lensless Imaging of Extended Scenes with Natural Illumination

In the astronomical applications for which coded aperture imaging was originally developed, the scene typically contains a discrete bright object in a largely dark field. A key assumption behind MURA encoding is that “all [nonzero] object points contribute a complete cyclic version of the basic [aperture] pattern,”<sup>3</sup> and the imaging geometry is designed to ensure that the complete aperture “shadows” are collected on the focal-plane array. However, terrestrial applications involve nonzero scene elements extending to and past the edge of the effective field-of-view: these scene elements will cast only partial aperture shadows on the focal plane. This effectively causes more object points to contribute to the image than there are image pixels, creating an inherently ill-posed problem for reconstructing the “true” scene.

To address the ill-posed imaging problem, we developed a new class of coded aperture masks which provide compelling advantages over MURAs. The new class of masks, which we call separable Doubly-Toeplitz apertures, provides

\*Address all correspondence to: Michael J. DeWeert, E-mail: [michael.deweert@baesystems.com](mailto:michael.deweert@baesystems.com)

images that are orders of magnitude faster to decode and are less affected by the ill-posed nature of the imaging of extended scenes—their transfer functions have far fewer zero eigenvalues than do those of MURA masks. Unlike MURAs, separable Doubly-Toeplitz apertures retain their properties when diffracted, making them much more robust for broadband imaging than MURAs.

This paper addresses the optics and mathematics of the new class of coded apertures.

The method has also been validated with real-world imaging of extended scenes in natural light. Such validations are absolutely necessary—much of the published work based on simulations alone presents overly optimistic predictions of coded-aperture performance. Thus, we provide a demonstration of coded-aperture imaging in visible light acquired with off-the-shelf hardware. This work also uncovered several optical phenomena that are minor nuisances for lensed imaging, but causes significant performance degradation for coded apertures, and would not have been uncovered by modeling and simulation alone. A future publication will provide in-depth exposition of the practical optical challenges for developing software-defined optics.

## 1.2 Former State-of-the-Art: Fourier Decoding and Modified Uniformly Redundant Array Masks

The coded-aperture mask is a flat plane parallel to an imaging array. In this configuration, the image  $\mathbf{I}$  is just a correlation function between the object  $\mathbf{O}$  and the mask  $\mathbf{M}$ :

$$\mathbf{I} = \mathbf{M} \otimes \mathbf{O}, \quad (1)$$

or in terms of matrix elements

$$I(i, j) \approx \sum_k \sum_l O(k, l) M(i + k, j + l). \quad (2)$$

The goal of decoding is to use knowledge of  $\mathbf{M}$  to recover the object scene  $\mathbf{O}$  from the encoded image  $\mathbf{I}$ .

The prior state-of-the-art in coded-aperture imaging was to select binary codes which could be inverted via forward correlation with slightly modified versions of themselves. Specifically, mask codes  $\mathbf{M}$  were selected for which the

masks are simply a collection of binary (0% or 100% transparent) elements and for which the inverse matrix  $\mathbf{G}$ , such that

$$\mathbf{O}_{\text{Estimated}} = \mathbf{G} \otimes \mathbf{I} = \mathbf{G} \otimes \mathbf{M} \otimes \mathbf{O} \approx \mathbf{O}, \quad (3)$$

was easily derivable from  $\mathbf{M}$  itself. For example, for URA<sup>3</sup> masks,

$$G_{i,j} = \begin{cases} +1, & M_{i,j} = 1 \\ -1, & M_{i,j} = 0 \end{cases}. \quad (4)$$

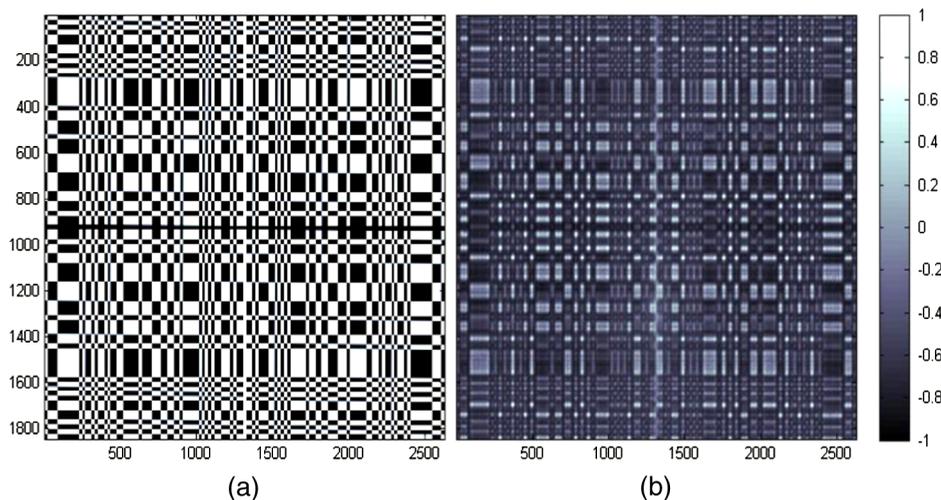
URAs and MURAs satisfy Eq. (3) under periodic boundary conditions. In practice, diffraction is not always negligible. An example of an ideal mask is shown in Fig. 1, along with a diffracted mask image. In this case, the MURA was lithographically deposited as a pattern on a glass slide, so that it was truly binary, either passing or blocking light. However, the collected mask image shows effects of scattering and diffraction, compromising the self-inverse property. In addition, as is explained below, even ideal MURAs have limited utility for extended scenes.

Because diffracted MURAs do not obey Eq. (3), inversion of the transfer function is required. The most straightforward approach for inversion of Eq. (1) is by Fourier transform. Our initial investigations using spatial-light-modulator (SLM)-defined masks showed noisy low-contrast recovered images. We successively eliminated the causes such as diffraction, polarization, and nonuniform SLM-element throughput. Finally, we traced the performance limitations to the ill-posed mathematics inherent in imaging extended scenes. The larger the mask, the more ill-posed the inversion, thus requiring regularization.

Initially, we implemented regularized decoding by introducing a real, positive function  $c^2(\mathbf{k})$  into the Fourier-transform-based decoding and computing an estimated image from

$$O_{\text{Recovered}}(\mathbf{k}) \approx I(\mathbf{k})M(\mathbf{k})/[|M(\mathbf{k})|^2 + c^2(\mathbf{k})]. \quad (5)$$

This is a special case of the Tikhonov<sup>5</sup> regularization method used in image deblurring<sup>6</sup>— $c^2(\mathbf{k})$  plays the dual



**Fig. 1** Example of a MURA: (a) ideal binary mask, and (b) image showing effects of diffraction. The collected mask image is no longer binary and exhibits numerous diffraction features.

roles of regularizing the inversion and controlling the effects of noise. With this regularization, recognizable decodings of laboratory images were achieved with both lithographically produced and programmable-SLM-based masks.

## 2 New Approach to Ill-Posed Coded-Aperture Decoding: Theory

Even with regularized Fourier inversion, decoded images of extended scenes were usually unrecognizable. We traced this to the Fourier method itself, which has the limitation of implicitly assuming cyclic or periodic behavior of the object  $\mathbf{O}$ . While the cyclic assumption is acceptable for compact objects against dark backgrounds, it is inappropriate for extended scenes viewed through large aperture masks.

To relax the assumption of cyclic boundaries, we recast the decoding problem in the most general terms. The correlation of a filter  $\mathbf{M}$  with an object scene  $\mathbf{O}$  to yield an image  $\mathbf{I}$  in Eq. (1) can also be expressed as a simple matrix multiplication:

$$\mathbf{I} = \mathbf{M}'' \cdot \mathbf{O}\mathbf{1}, \quad \text{so that } I_i = \sum_k M''_{i,k} O_{1k}, \quad (6)$$

where  $\mathbf{I}$  and  $\mathbf{O}\mathbf{1}$  are the one-dimensional (1-D) vectors created by concatenating the rows of  $\mathbf{I}$  and  $\mathbf{O}$ :

$$I_{C_I \times i+j} = I_{i,j}, \quad \text{and} \quad O_{1, C_O \times i+j} = O_{i,j}, \quad (7)$$

where  $C_I$  and  $C_O$  are the numbers of columns in  $\mathbf{I}$  and  $\mathbf{O}$ , respectively. The matrix  $\mathbf{M}''$  has rows that are shifted copies of the first row, which is constructed from the filter  $\mathbf{M}$  by embedding each row of  $\mathbf{M}$  into a segment of length  $C_O$ , padded with zeros for elements  $C_F + 1$  through  $C_O$ , then concatenating the segments into a row of length  $C_O \times R_O$ , where  $R_O$  is the number of rows in  $\mathbf{O}$ . That is,

$$M''_{1, C_O \times i+j} = \begin{cases} M_{i,j} & \text{for } j \in \text{the first } C_F \text{ columns} \\ 0 & \text{for } j \in \text{later columns} \end{cases}. \quad (8)$$

The size of the matrix  $\mathbf{M}''$  grows as the sizes of the image and the filter mask grow. If the image has dimensions  $r_I \times c_I$  and the filter mask has dimensions  $r_F \times c_F$ , then the sampled object is of size  $r_O \times c_O$ , with  $r_O = (r_I + r_F - 1)$  and  $c_O = (c_I + c_F - 1)$ . This yields for  $\mathbf{M}''$  dimensions of  $(r_I \times c_I)$  rows and  $(r_O \times c_O)$  columns.

The relationship between image size, filter size, and the extent of the object area sampled illustrates the inherently ill-posed nature of any imaging system. Since the PSF of a real-optical system is never a delta-function, there are always more object points contributing to an image than there are points in the image itself. For conventional imaging systems, where the PSF may be very compact, it is possible to clip or “throw-away” points near a reconstructed image’s edges and to retain only those which represent an approximately well-posed inversion. If significant diffraction is present, then the clipping strategy is inadequate—hence, the emphasis is on the diffraction limit in conventional optics.

For coded-aperture imaging, the filter area (and the spread function) is directly proportional to the light-gathering power of the system. Because of this proportionality of spread to light-gathering, it is not possible to use a clipping strategy to select well-posed portions of the recovered image, except in the trivial case of a single-pinhole aperture.

Thus coded-aperture imaging must address inherently ill-posed encoding and decoding.

While the forward correlation can be done exactly using Eq. (6), the inversion (estimating  $\mathbf{O}$  from the measurements  $\mathbf{I}$ ) presents a seriously ill-posed<sup>7</sup> and computationally challenging task. The general case is commonly solved via the regularized normal equations:

$$\mathbf{M}''^T \mathbf{I} = (\mathbf{M}''^T \mathbf{M}'' + \alpha^2 \mathbf{1}_{r_O \times c_O}) \mathbf{O}\mathbf{1}, \quad (9)$$

where  $\mathbf{1}_{r_O \times c_O}$  is the identity matrix with each dimension measuring

$$r_O \times c_O = (r_I + r_F - 1) \times (c_I + c_F - 1). \quad (10)$$

The operations count for Eq. (9) is on the order of<sup>8</sup>

$$N_{\text{Operations, Generalized}} \propto (r_O c_O)^3. \quad (11)$$

For 1 megapixel images, this yields  $10^{18}$  operations, requiring on the order of  $10^6$  s to 12 days on a teraflops processor. Thus, computations on large images using generalized masks are usually impractical. Further, the matrix  $\mathbf{M}''$  is so large that the accumulation of round-off errors will likely swamp the inversion even if it could be accomplished in a manageable time. The only exception would be if the encoding mask is a single-pixel pinhole aperture—a well-posed but trivial case.

### 2.1 Separable Doubly-Toeplitz Masks

We have developed a method to specify finite-sized masks that greatly reduce the degree to which the imaging is ill-posed and render the inversions of Eq. (9) tractable. If the mask  $\mathbf{M}$  can be expressed as an outer product of two 1-D vectors  $\mathbf{A}$  and  $\mathbf{B}$ , with lengths  $N_A$  and  $N_B$ , respectively, then  $\mathbf{M}$  is an  $N_A \times N_B$ -dimensional matrix of the form

$$M(i, j) = A(i)B(j), \quad (12)$$

where  $A(i)$  is the  $i$ 'th element of vector  $\mathbf{A}$ , and  $B(j)$  is the  $j$ 'th element of vector  $\mathbf{B}$ . Equation (9) can now be written simply as a product of much smaller two-dimensional (2-D) matrices:

$$\mathbf{I} = \mathbf{M}_A \mathbf{O} \mathbf{M}_B^T. \quad (13)$$

The matrices  $\mathbf{M}_A$  and  $\mathbf{M}_B$  are Toeplitz, with the elements shifted by 1 pixel on each successive row:

$$\mathbf{M}_A = \begin{bmatrix} A_1 & A_2 & \dots & A_{N_A} & 0 & 0 & \dots & 0 \\ 0 & A_1 & A_2 & \dots & A_{N_A} & 0 & 0 & \dots & 0 \\ \vdots & \ddots \end{bmatrix}. \quad (14)$$

The vector  $\mathbf{M}_B$  is similarly constructed from the elements of the vector  $\mathbf{B}$ .

The difficulty of inversion using Eq. (13) is vastly less than for solving Eq. (9). While this is still an ill-posed problem, the sizes of matrices  $\mathbf{M}_A$  and  $\mathbf{M}_B$  are much smaller than  $\mathbf{M}''$  in Eq. (9). With the matrix dimensions summarized in

**Table 1** Relationship of dimensions of the collected image, the encoding filter mask, and the object area contributing to the image.

Matrix	Rows	Columns
Coded-aperture mask filter, $\mathbf{M}$	$r_F$	$c_F$
Image on focal plane, $\mathbf{I}$	$r_I$	$c_I$
Object area, $\mathbf{O}$	$r_O = (r_I + r_F - 1)$	$c_O = (c_I + c_F - 1)$
Left Toeplitz mask, $\mathbf{M}_A$	$r_I$	$r_O$
Right Toeplitz mask, $\mathbf{M}_B$	$c_I$	$c_O$

Table 1, the inversions require an operations count on the order of

$$N_{\text{Operations, Doubly-Toeplitz}} \propto [(r_I r_O)^3 + (c_I c_O)^3]. \quad (15)$$

For image sizes on the order of 1 megapixel, this yields on the order of  $2 \times 10^9$  operations, requiring just milliseconds on a teraflops processor. Thus, unlike arbitrary general masks or diffracted MURAs, separable Doubly-Toeplitz masks can be inverted in a usefully short time.

To our knowledge, this Doubly-Toeplitz strategy for specifying coded apertures has not been previously applied to the spatial imaging problem, though there are hints of 1-D Toeplitz masks in the temporal motion-deblurring work of Agrawal and Raskar.<sup>9,10</sup>

### 3 Diffraction Effects on Doubly-Toeplitz Masks

Having identified an approach to reduce the impact of ill-posed inversions, the investigations next turned to mitigating the effects of diffraction. Simulations verified that the Doubly-Toeplitz property is preserved in a diffracted mask. Simulations also provide a means of synthesizing decoding matrices representing diffracted mask images for use when good laboratory-collected mask images are not available.

Diffraction simulations are particularly straightforward for a collection of identical rectangular apertures placed at points on a regular-rectangular-grid SLM. Let  $\mathbf{I}_O$  and  $\mathbf{I}_P$  denote the intensity fields at the object and image planes, respectively, and let  $\mathbf{E}_O$  denote the complex optical field at the object plane. When the illumination is incoherent, the diffracted intensity in a pixel at location  $(i, j)$  on the image plane is

$$I_P(i, j) \approx \sum_{k,l} |Q_X(i+k)|^2 |I_O(k, l)| |Q_Y^T(l+j)|^2, \quad (16)$$

where

$$Q_X(k_x) = \sum_{j_x} A(j_x) \exp \left\{ i \frac{2\pi}{\lambda} j_x p_X \left[ k_x \frac{\Delta x_P}{f_P} - \frac{j_x p_X}{2} \left( \frac{1}{R} + \frac{1}{f_P} \right) \right] \right\} \times \text{sinc} \left\{ \frac{\pi}{\lambda} a_X \left[ k_x \frac{\Delta x_P}{f_P} - j_x p_X \left( \frac{1}{R} + \frac{1}{f_P} \right) \right] \right\} \quad (17)$$

and

$$Q_Y(k_y) = \sum_{j_y} B(j_y) \exp \left\{ i \frac{2\pi}{\lambda} j_y p_Y \left[ k_y \frac{\Delta y_P}{f_P} - \frac{j_y p_Y}{2} \left( \frac{1}{R} + \frac{1}{f_P} \right) \right] \right\} \times \text{sinc} \left\{ \frac{\pi}{\lambda} a_Y \left[ k_y \frac{\Delta y_P}{f_P} - j_y p_Y \left( \frac{1}{R} + \frac{1}{f_P} \right) \right] \right\}. \quad (18)$$

The parameters in Eqs. (16)–(18) are

$(a_X, a_Y)$  = aperture opening sizes,

$(p_X, p_Y)$  = aperture pitch,  $> (a_X, a_Y)$ ,

$(\Delta x_P, \Delta y_P)$  = image-plane pixel pitch,

$\mathbf{A}, \mathbf{B}$  = vectors of weights of mask elements in the  $x$  and  $y$  directions,

$R$  = range from aperture mask to object plane,

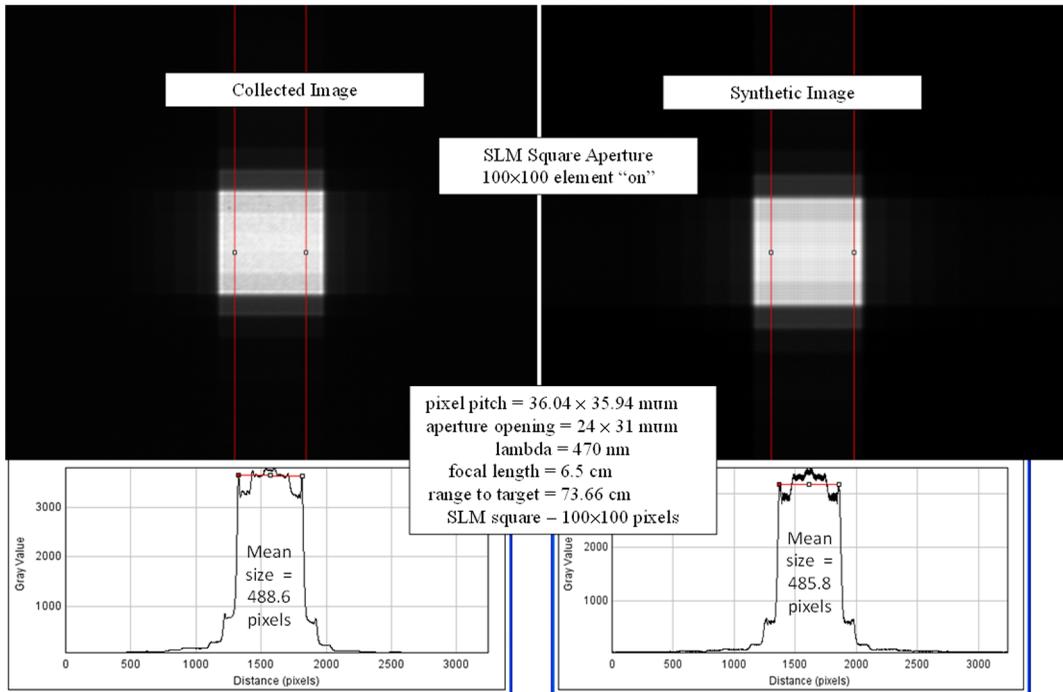
$f_P$  = range from aperture mask to image plane, and

$\lambda$  = wavelength of light.

Diffraction is a significant contributor to extended-scene coded-aperture imaging in visible light. Our experimental investigations were conducted with a bandpass filter centered on 550 nm, and a coded aperture implemented by varying the throughput of elements of a transmissive SLM. The SLM had elements with dimensions of approximately  $24 \mu\text{m} \times 31 \mu\text{m}$  on a  $36\text{-}\mu\text{m}$  pitch. The total area of active pixels measured  $1024 \times 768$  pixels, corresponding to a maximum aperture of approximately  $27 \text{ mm} \times 36 \text{ mm}$ . The distance from the aperture to the focal plane was typically 65 mm. With these parameter values, the Fresnel number for the coded aperture optics ranged from under 0.02 (with a single SLM element used as a pinhole) to over 40,000. Because the objective of coded-aperture imaging is to capture the advantages of pinhole imaging while maximizing light-gathering capability, many SLM elements were set to the “on” state, so that our experiments were conducted predominantly at Fresnel numbers between 150 and 40,000. We simulated the Fresnel diffraction with a custom-written numerical code developed for coded-aperture imaging and validated by comparison with laboratory results using both programmable and fixed apertures. Results of one of these experiments are shown in Fig. 2. Within the experimental uncertainties caused by random variations of the transmissivity of the SLM elements and the finite resolution of the focal plane, the agreement is very good. In future papers, we plan to present the effects of diffraction, noise, and variable transmissivity on the quality of coded-aperture images via objective and subjective metrics such as peak signal-to-noise ratio and structural similarity.

Examination of Eq. (16) shows that, if the raw aperture mask is an outer product of two 1-D vectors  $\mathbf{A}$  and  $\mathbf{B}$ , then the effects of the diffracted mask can still be expressed as an operation of a pair of Toeplitz matrices. Thus, a separable Doubly-Toeplitz aperture mask will yield a separable Doubly-Toeplitz encoding even when there is significant diffraction. If the object is an incoherent point-like source, then its image will be decomposable into  $\mathbf{A}$  and  $\mathbf{B}$  vectors which can be used to synthesize decoding masks.

To verify the diffraction model, we conducted a series of experiments comparing simulations with image data, obtaining excellent agreement between model and data. Diffracted coded apertures were also modeled and compared with “collected masks”—images of a point-source viewed



**Fig. 2** Images of a point source viewed through a single-large square aperture synthesized by setting a 100-pixel-square area of SLM pixels to the “on” state. Vertical cuts through the images illustrate the diffraction features. The agreement is very good, despite noise and the element-to-element variability of the SLM.

through the spatial-coding masks. We verified that the collected masks decompose into the outer products of diffracted 1-D functions, in accord with Eq. (16). The mask shown in Fig. 3 is the basic “workhorse” mask used in the investigations presented in this paper. Three scaled versions of this mask were used: “02” is 2.2-mm wide, “04” is 4.4-mm wide, and “08” is 8.8-mm wide. The “02” mask spans 62 elements of our SLM, with a minimum feature size of two SLM elements. The SLM elements had a nominal 35.5- $\mu\text{m}$  pitch. The pitches were identical along the axes of the SLM, but the aperture openings were asymmetric—24- $\mu\text{m}$  high and 31- $\mu\text{m}$  wide, accounting for the asymmetry of the diffraction pattern. The 31-element base mask<sup>10</sup> for temporal encoding and decoding to compensate for motion blur. In a future publication, we will report on methods for efficient searches for masks optimized specifically for 2-D spatial coded-aperture imaging.

Figure 3 also demonstrates that the diffraction increases the size of the effective encoding masks, increasing the ill-posedness of the inverse problem. This makes separable Doubly-Toeplitz masks even more valuable for cases in which fine imaging resolution is required.

#### 4 Solution: Method

In real-imaging systems, noise will always arise from effects such as photon counting or read noise, so that the actual inverse problem includes a noise term  $\mathbf{N}$ . The matrix  $\mathbf{I}^\delta$  denotes the image containing noise effects which are not encoded by the aperture mask. In addition, there is the ill-posed nature of the problem embodied in the nonsquare matrices  $\mathbf{M}_A$  and  $\mathbf{M}_B$ . Both the noise and the ill-posedness

require a regularization method. We begin by multiplying Eq. (13) by the transposes of the mask matrices, obtaining the normal equation:

$$\mathbf{M}_A^T \mathbf{I}^\delta \mathbf{M}_B = (\mathbf{M}_A^T \mathbf{M}_A) \mathbf{O} (\mathbf{M}_B^T \mathbf{M}_B)^T. \quad (19)$$

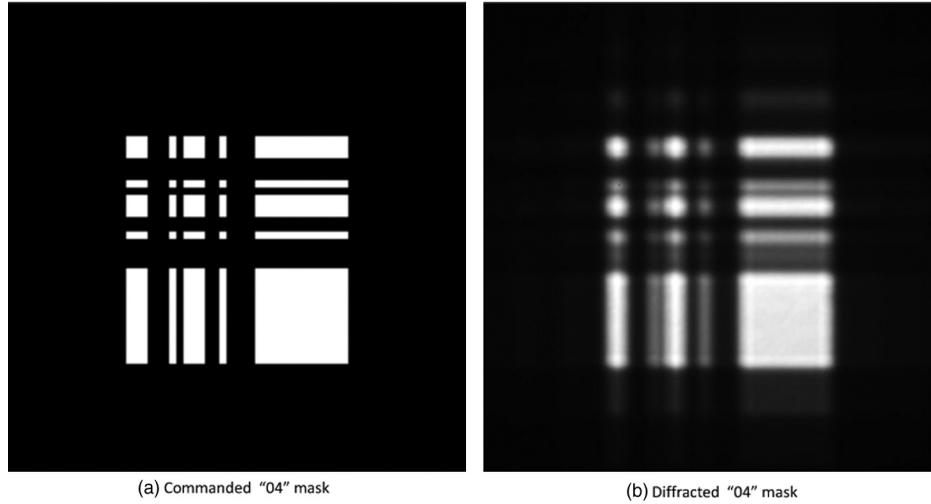
The matrix products  $(\mathbf{M}_A^T \mathbf{M}_A)$  and  $(\mathbf{M}_B^T \mathbf{M}_B)$  are square and symmetric, but are not invertible, due to degenerated zero eigenvalues. The degree of degeneracy increases with the size of the filter masks—there are  $(r_o - r_l)$  zero eigenvalues for  $\mathbf{M}_A$ , and  $(c_o - c_l)$  zero eigenvalues for  $\mathbf{M}_B$ . [Compare this with solving Eq. (9) for a MURA or generic mask, which is orders of magnitude more degenerate, having  $(r_o c_o - r_l c_l)$  zero eigenvalues.]

To relieve the degeneracies, we introduce regularization parameters  $\alpha_A$  and  $\alpha_B$ , so that

$$\mathbf{M}_A^T \mathbf{I} \mathbf{M}_B \approx (\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2 \mathbf{1}_{R_o}) \mathbf{O} (\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2 \mathbf{1}_{C_o})^T + \mathbf{M}_A^T \mathbf{N} \mathbf{M}_B, \quad (20)$$

where  $\mathbf{1}_{R_o}$  is an identity matrix with dimensions  $R_o$ , and  $\mathbf{1}_{C_o}$  is an identity matrix with dimensions  $C_o$ . Equation (20) now multiplies  $\mathbf{O}$  by invertible matrices, so that a noisy estimate of  $\mathbf{O}$  can be derived from the image:

$$\mathbf{O}_{\text{Est}} = (\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2 \mathbf{1}_{R_o})^{-1} \mathbf{M}_A^T \mathbf{I} \mathbf{M}_B \times (\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2 \mathbf{1}_{C_o})^{-1}. \quad (21)$$



**Fig. 3** (a) Binary mask commanded on the SLM, and (b) the image of a point source viewed through the mask, showing Fresnel diffraction features.

## 5 Methods for Improving the Solution

### 5.1 Normalization of Reconstructed Images

The reconstructed images initially show patterns at their edges within a distance of one mask width (or height) of the border. We noted similar artifacts in the encoded images of uniform (all white) scenes, which led us to a method by which the artifacts could be normalized out of the reconstructed images. The normalization is accomplished by encoding and decoding a uniform all-white object  $\mathbf{W}$  to create a matrix  $\mathbf{W}_{\text{Est}}$  that includes the spurious brightness differences accrued as the coded-aperture mask straddles the edge of the image area:

$$\mathbf{W}_{\text{Est}} = [\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2 \mathbf{1}_{R_o \times R_o}]^{-1} (\mathbf{M}_A^T \mathbf{M}_A) \mathbf{W} (\mathbf{M}_B^T \mathbf{M}_B) \times [\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2 \mathbf{1}_{C_o \times C_o}]^{-1}. \quad (22)$$

$\mathbf{O}_{\text{Est}}$  is divided by  $\mathbf{W}_{\text{Est}}$ , element-by-element, to obtain a normalized object estimate  $\mathbf{O}_{\text{Norm}}$ :

$$\mathbf{O}_{\text{Norm}}(i, j) = \mathbf{O}_{\text{Est}}(i, j) / \mathbf{W}_{\text{Est}}(i, j). \quad (23)$$

The normalization can also be expressed as multiplication by two normalization matrices  $\mathbf{D}_A$  and  $\mathbf{D}_B$ :

$$\mathbf{O}_{\text{Norm}} = \mathbf{D}_A \mathbf{O}_{\text{Est}} \mathbf{D}_B^T. \quad (24)$$

This is because the all-white matrix  $\mathbf{W}$  can be expressed as an outer product of two 1-D vectors of constant value, which in turn allows the normalization estimate to be decomposed into an outer product of two 1-D matrices:

$$\mathbf{W}_{\text{Est}}(i, j) = W_A(i) W_B(j). \quad (25)$$

The normalization matrices  $\mathbf{D}_A$  and  $\mathbf{D}_B$  each have identical columns, given by

$$D_A(i, j) = 1/W_A(i) = 1 / \left\{ \sum_{k=1}^{R_o} [(\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2 \mathbf{1}_{R_o \times R_o})^{-1} \times (\mathbf{M}_A^T \mathbf{M}_A)](i, k) \right\}$$

$$D_B(i, j) = 1/W_B(i) = 1 / \left\{ \sum_{k=1}^{C_o} [(\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2 \mathbf{1}_{C_o \times C_o})^{-1} \times (\mathbf{M}_B^T \mathbf{M}_B)](i, k) \right\}. \quad (26)$$

The decoded and normalized object estimates are then

$$\mathbf{O}_{\text{Norm}} = \mathbf{D}_A (\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2)^{-1} \mathbf{M}_A^T \mathbf{I}^\delta \mathbf{M}_B \times [(\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2)^T]^{-1} \mathbf{D}_B^T. \quad (27)$$

### 5.2 Improvement via Landweber Iteration

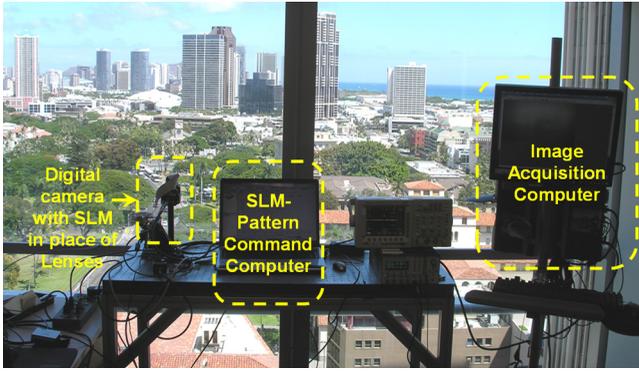
The results can be further improved by iterating<sup>11</sup> from the solution in Eq. (27). We settled on the method of Landweber,<sup>12</sup> which iterates using a regularization constant  $\varepsilon_L$ :

$$\mathbf{O}_{(n+1)} = \mathbf{O}_{(n)} + \varepsilon_L \mathbf{M}_A^T \mathbf{P}_A [\mathbf{I}^\delta - \mathbf{M}_A \mathbf{O}_{(n)} \mathbf{M}_B^T] \mathbf{P}_B^T \mathbf{M}_B, \quad (28)$$

where  $\mathbf{P}_A$  and  $\mathbf{P}_B$  are the positive-definite symmetric matrices specified to control the convergence. While classic Landweber implementation uses  $\mathbf{P}_A = \mathbf{P}_B = \mathbf{1}$  (the identity matrix), our implementation uses

$$\mathbf{P}_A = \mathbf{D}_A \mathbf{S}_A \quad \text{and} \quad \mathbf{P}_B = \mathbf{D}_B \mathbf{S}_B, \quad (29)$$

where  $\mathbf{S}_A$  and  $\mathbf{S}_B$  are the inverses:



**Fig. 4** Equipment for imaging outdoor scenes. This view looks from downtown Honolulu eastward toward Diamondhead mountain.

$$\begin{aligned} \mathbf{S}_A &= (\mathbf{M}_A^T \mathbf{M}_A + \alpha_A^2)^{-1}, \quad \text{and} \\ \mathbf{S}_B &= (\mathbf{M}_B^T \mathbf{M}_B + \alpha_B^2)^{-1}. \end{aligned} \quad (30)$$

The starting estimate  $\mathbf{O}_{(0)}$  is estimated from Eq. (27), then iterated using

$$\begin{aligned} \mathbf{O}_{(n+1)} &= \mathbf{O}_{(n)} + \varepsilon_L^2 \mathbf{D}_A \mathbf{S}_A \mathbf{M}_A^T [\mathbf{I}^\delta - \mathbf{M}_A \mathbf{O}_{(n)} \mathbf{M}_B^T] \mathbf{M}_B \mathbf{S}_B \mathbf{D}_B \\ &\quad + a_n (\mathbf{O}_{(n)} - \mathbf{O}_{(0)}), \end{aligned} \quad (31)$$

with

$$a_n = 1/(4 + n). \quad (32)$$

Equation (31) is actually a modified Landweber method<sup>13</sup> that is guaranteed to converge as long as  $0 < \varepsilon_L^2 < 2$ . However, the convergent solution is not guaranteed to be correct—the results may be noise dominated unless the iterations stop at a suitable point. We apply an *a posteriori* criterion based on the residuals after  $n$  iterations:

$$\mathbf{R}_{(n)} = \mathbf{M}_A^T [\mathbf{I}^\delta - \mathbf{M}_A \mathbf{O}_{(n)} \mathbf{M}_B^T] \mathbf{M}_B, \quad (33)$$

with norm

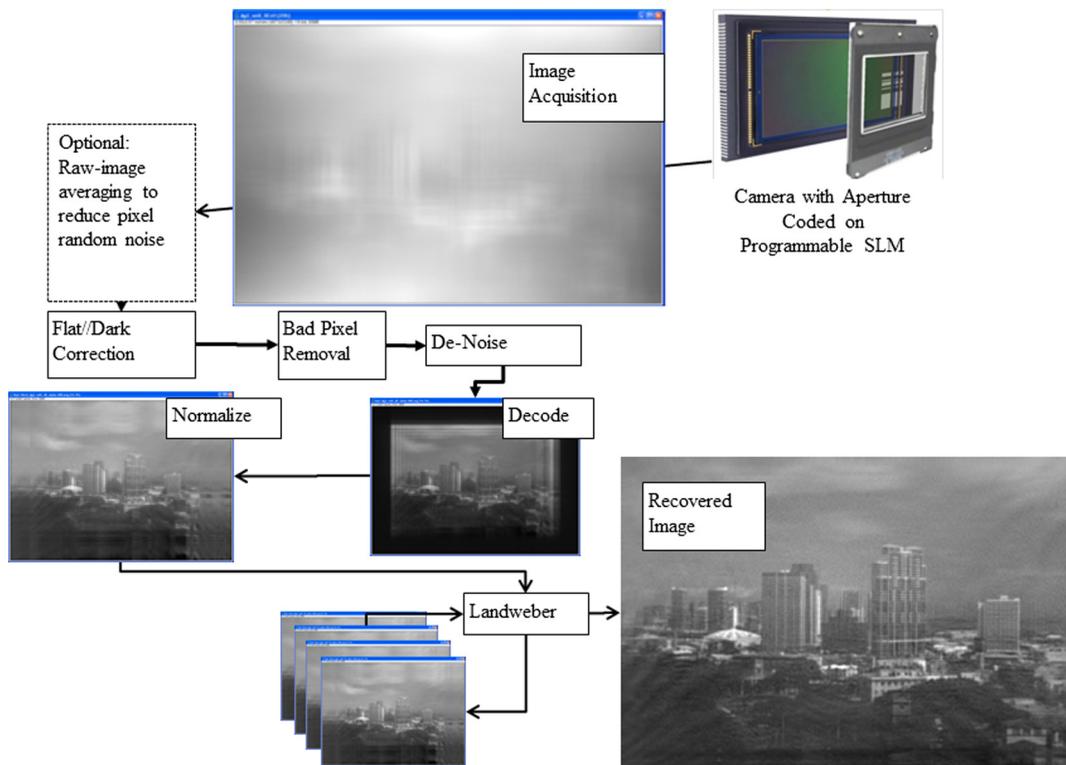
$$\|r_{(n)}\|^2 = \sum_{i,j} [\mathbf{R}_{(n)}(i,j)]^2. \quad (34)$$

The algorithm stops the iterations when the residuals from Eq. (34) begin to increase.

## 6 Encoding and Decoding Real-Extended Imageries

The refined methods of coding and decoding made it possible to image extended outdoor scenes in natural light. We used these images to further improve and debug the processing.

One experimental setup is shown in Fig. 4. The camera was situated in a commercial high-rise building in Honolulu, Hawaii, approximately 45 m above the ground, looking east toward the city skyline. The SLM was a Holoeye HEO 0017 LC-SLM with a  $1024 \times 768$ -element active area. This was paired with a Kodak RMV 16000 digital camera.



**Fig. 5** Processing flow for coded-aperture image acquisition and decoding. Raw coded-aperture imagery is unrecognizable—processing yields a clear image. Data conditioning reduces numerous sources of unencoded noise and interference, so that the decoding algorithms can compute a recognizable image.

The RMV 16000 was chosen because its focal plane area of  $36 \text{ mm} \times 24 \text{ mm}$  is close in size to the  $33 \text{ mm} \times 25 \text{ mm}$  dimensions of the Holoeye SLM and because the RMV pixel pitch of  $7.4 \mu\text{m}$  allows oversampling of the nominal  $35.5\text{-}\mu\text{m}$  pitch of the SLM elements, reducing sampling as a source of error. The SLM was placed  $6.5 \text{ cm}$  in front of the camera's focal plane.

In addition to the programmable SLM, the camera had a green-bandpass filter to reject optical wavelengths that are



(a) 02 Mask 0-deg Mask Orientation



(b) 02 Mask Four Mask Orientation Summed



(c)

**Fig. 6** Final results (b) of processing real imagery with multiple masks (four 90-deg rotations of the “02” mask), versus using a single mask (a), showing reduction of decoding “ghost” artifacts. (c) The reference photo of the scene acquired with a conventional camera. (The clouds are different because the images were acquired at different times.)

not adequately blocked by the “off” pixels of the SLM. We also developed methods to deal with “bad” or “dead” pixels, inhomogeneous focal plane response, fixed-pattern noise, and with stray light reflecting inside the camera. Most of the effects are minor nuisances for conventional lens-assisted imaging, but cause significant degradation of coded-aperture decoding quality. This is because the decoding algorithms attempt to invert the image, artifacts and all, to construct an object scene. Careful reduction of all noncoded-light is required to achieve satisfactory results.

### 6.1 Encoding and Decoding Real-Extended Imagery: Results

The details of the various artifacts and the associated mitigation algorithms will be presented in a future publication. The basic steps to precondition the images then to decode them are outlined in Fig. 5. The denoising step can be either single-frame denoising or denoising by adding multiple raw images before decoding. Stray-light interference limited the effectiveness of single-frame denoising. This was mitigated via a 20-frame average, with the raw images averaged together before application of flat, dark, and defect corrections.

### 6.2 Multimask Image Stacking

Our investigations showed that the regularization artifacts differ from mask to mask. Thus, reconstruction could be improved by using multiple images collected with a sequence of different masks. We tested this approach using four 90-deg rotations of the same basic mask. The four images were decoded separately and stacked to provide an improved image. Results from simply averaging the four resulting decoded images are compared with single-mask imagery in Fig. 6 and also compared with an image acquired with a conventional lensed camera.

## 7 Conclusions

Our goal was to improve the performance of completely lensless imagers, especially for imaging extended scenes in natural light. We have demonstrated a novel approach, separable Doubly-Toeplitz masks, for specifying coded-aperture masks that allow efficient decoding and have shown the feasibility of using programmable SLMs to implement the new class of masks. This work presents a fundamental change in coded-aperture imaging and lays the foundation for future developments in which we will quantify the performance of the separable Doubly-Toeplitz masks for a greater variety of scenes with both programmable and fixed masks, exploring the limitations of the undistorted field-of-view achievable, as well as extending the applicability to bands beyond the visible spectrum.

In particular, it has been suggested that the diffraction could limit the performance of coded apertures for infrared imaging.<sup>14</sup> The work presented here shows that the diffraction can be readily included in the decoding algorithms for separable Doubly-Toeplitz masks. More serious limiters, especially for the long-wave infrared, are likely to be the influence of nonencoded information from thermal self-emission of the camera body and drifts in the nonuniformity corrections for uncooled thermal-camera focal planes.

Future work will also explore how the separable-Doubly-Toeplitz masks can be combined with refractive elements, extending the super-resolution capabilities studied under the DARPA LACOSTE program.<sup>15</sup>

The project has also identified other implementation-specific difficulties that need to be addressed in using SLMs for lensless imaging. Since it is still early in the development of the technology, rapid progress can be made. Once the device-specific sources of noise and error are mitigated, the next biggest improvement will come from developing a computationally efficient means for identifying mask candidates. Rather than generate random binary codes, then exhaustively test each one, a more-efficient approach is likely to be a genetic algorithm which evolves mask codes from the best currently known starting points.

Since Doubly-Toeplitz masks are continuous, having no isolated open sections, self-supporting nonprogrammable masks can be built in to the fuselages and carapaces of small unmanned vehicles with three-dimensional printing methods, reducing the time and cost of fielding new systems. If fabricated from material that blocks x-ray and gamma-ray radiations, then this would in turn enable extremely compact noncontact imaging in these wavelengths for medicine, defense, and homeland security.

#### Acknowledgments

We are grateful for support for this work provided by the Office of Naval Research under contract N00014-10-C-0193.

#### References

1. R. H. Dicke, "Scatter-hole cameras for x-rays and gamma rays," *J. Astrophys.* **153**, L101–L106 (1968).
2. J. G. Ables, "Fourier transform photography: a new method for x-ray astronomy," *Proc. Astron. Soc. Aust.* **1**(4), 172–173 (1968).
3. E. E. Fenimore and T. M. Canon, "Coded aperture imaging with uniformly redundant arrays," *Appl. Opt.* **17**(3), 337–347 (1978).
4. S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Appl. Opt.* **28**(20), 4344–4352 (1989).
5. A. N. Tikhonov et al., *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer Academic Publishers, Moscow, Russia (1995).
6. S. Kimitei, *Algorithms for Toeplitz Matrices and Their Applications to Image Reconstruction*, Georgia State University, Atlanta, Georgia (2005).
7. J. Hadamard, "Sur les problèmes aux dérivées partielles et leur signification physique," *Princeton Univ. Bull.* **13**(4), 49–52 (1902).
8. W. H. Press et al., Eds., "Highly-optimized algorithms may require just  $N(\log_2 7)$  operation instead of  $N^3$  operations," in *Numerical Recipes in C—The Art of Scientific Computing*, 2nd ed., pp. 102–104, Cambridge University Press, New York (1999).
9. R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: motion deblurring using fluttered shutter," *ACM Trans. Graphics (TOG) Proc. ACM SIGGRAPH* **24**(3), 795–804 (2006).
10. A. Agrawal and R. Raskar, "Resolving objects at higher resolution from a single motion-blurred image," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition 'CVPR'*, pp. 1–8, IEEE Computer Society, Minneapolis, Minnesota (2007).
11. W. H. Press et al., *Numerical Recipes in C—The Art of Scientific Computing*, 2nd ed., pp. 55–58, Cambridge University Press, New York (1999).
12. T. Elfving and T. Nikazad, "Stopping rules for Landweber-type iteration," *Inverse Probl.* **23**, 1417–1432 (2007).
13. H. Egger, "On the convergence of modified Landweber iterations for nonlinear inverse problems," SFB Report No. 2010-017, Karl-Franzens Universität Graz (2010).
14. S. R. Gottesman, "Coded apertures: past, present, and future application and design," *Proc. SPIE* **6714**, 671405 (2007).
15. A. Mahalanobis et al., "IR performance study of an adaptive coded aperture "diffractive imaging" system employing MEMS "eyelid shutter" technologies," *Proc. SPIE* **6714**, 67140D (2007).

**Michael J. DeWeert** is a BAE Systems Engineering Fellow and the chief technology officer at BAE Systems in Honolulu. He has over 25 years of experience in remote sensing research and development for intelligence, surveillance, and reconnaissance. His current responsibilities include technical oversight spectral imaging, LIDAR, interferometric remote-sensing systems, and modeling and simulation for maritime and terrestrial sensing. Currently, he is the principal investigator for projects developing advanced imaging technologies for long-standoff mine countermeasures.

**Brian P. Farm** has over 25 years of experience in modeling, algorithm/software development, and spatial and image data analyses. He is involved in algorithm/software research and development for data exploitation, real-time image analysis, and sensor-model adjustment and exploitation. He has worked on 3-D imaging LIDARs, developed registration and processing algorithms for imaging and interferometric systems, and developed information processing for novel, nontraditional imaging sensors, and created simulation systems to allow advanced algorithm testing on proposed new sensor systems.