

Multichannel sampling schemes for optical imaging systems

Andrew D. Portnoy,^{1,*} Nikos P. Pitsianis,^{1,2} Xiaobai Sun,² and David J. Brady¹

¹Fitzpatrick Institute for Photonics and Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708, USA

²Department of Computer Science, Duke University, Durham, North Carolina 27708, USA

*Corresponding author: adp4@duke.edu

Received 10 September 2007; revised 4 January 2008; accepted 11 January 2008;
posted 24 January 2008 (Doc. ID 87029); published 19 March 2008

We introduce a framework of focal-plane coding schemes for multichannel sampling in optical systems. A particular objective is to develop an ultrathin imager without compromising image resolution. We present a complete $f/2.1$ optical system with a thickness of 2.2 mm. The resolution is maintained in the thin optical system by an integrated design of the encoding scheme, the process of making the coding elements, and the decoding algorithms. © 2008 Optical Society of America

OCIS codes: 100.3190, 110.1758, 110.3010.

1. Introduction

The generalized sampling theory by Papoulis [1] has been applied in multiband or multichannel imaging [2–4]. In particular, several research groups have focused on the application of the theory to “superresolved” optical imaging [5–8]. Previous implementations of multichannel sampling strategies have primarily utilized optical techniques to optimize in certain specific system metrics. For example, multichannel sampling with multiple aperture lenslet arrays has been used in the TOMBO system [9] to substantially reduce imaging system thickness and volume. Broadly, there is considerable interest in improving or optimizing performance metrics for computational imagers based on multichannel sampling. The Compressive Optical MONTAGE Photography Initiative (COMP-I) [10–12] has explored new strategies for multichannel imagers by a codesign of the optics, the electronic sampling strategy, and computational reconstruction scheme [13].

We describe the formal basis of the focal-plane coding strategies. We illustrate this in particular with an analysis and comparison of the TOMBO scheme

and a multiplexing scheme developed for the COMP-I system [11]. Both systems utilize thin imaging optics using a lenslet array, multiapertures, and computational integration of a single higher-resolution image from multiple images. They differ, however, not only in physical implementations—one is by lenslet displacements and the other by focal-plane masks—but also in the complexity and stability of computational reconstruction procedures. As the computation reconstruction becomes indispensable in such systems, it is an important system design issue to assess both the computational efficiency and the additional errors arising from the numerical procedures. We also describe the COMP-I implementation of the coding schemes with lenslet masks and present experimental results.

The rest of the paper is organized as follows. We introduce in Section 2 the focal-plane coding framework. In Section 3 we detail the implementation of the system components that support the focal-plane coding. We present in Section 4 experimental results, and in Section 5 we present numerical image reconstruction. We conclude the paper in Section 6.

2. Focal-Plane Coding: Model, Systems, and Analysis

Multichannel sampling is well understood in concept since the seminal work by Papoulis [1]. We introduce

a novel realization of the multichannel sampling theory on an optical system in order to reduce camera thickness without compromising image resolution. We present a framework of multichannel sampling via a lenslet array as well as the algebraic procedures for decoding, i.e., constructing a single image without the loss of resolution.

We illustrate the principle and practice with two particular families of focal-plane coding schemes. One is implemented with lenslet displacements, which was first introduced and demonstrated in the TOMBO system [9], and the other is implemented with coded masks at the lenslets. Both are implemented in the COMP-I system [14]. The two coding schemes also have different numerical behaviors in computational image reconstruction.

We first present a unifying model of optical imaging systems with lenslet arrays.

A. System Model per Aperture

In a typical image system, the object field, $f(x, y)$, is blurred by a point spread function (PSF), which is a characteristic of the imaging system. In an ideal case, the PSF is shift invariant and can be represented by a bivariate function $h(x, y)$. The blurred image is electronically sampled by a two-dimensional detector array, $G = [g_{ij}]$, where g_{ij} is the measurement at pixel (i, j) . We describe the extent of the focal plane by $[-X/2, X/2] \times [-Y/2, Y/2]$. In the case of incoherent imaging, the transformation from the source intensity $f(x, y)$ in object space to the discrete data array $G = [g_{ij}]$ measured by the detector may be modeled as follows:

$$g_{ij} = \int_{-X/2}^{X/2} \int_{-Y/2}^{Y/2} s_{ij}(x, y, \tilde{f}(x, y)) dx dy, \quad (1)$$

$$\tilde{f}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) h(\alpha x - \xi, \alpha y - \eta) d\xi d\eta,$$

where \tilde{f} is the blurred and scaled image at the focal plane, modeled by the convolution of the object function with the PSF, α is a system-dependent scaling parameter, as we shall illustrate with some cases shortly, and the function s_{ij} characterizes the sampling at the (i, j) pixel of the blurred image at the focal plane. The support of $s_{ij}(x, y)$ may be limited to the geometry and location of the pixel at the detector.

All the pixels at the detector assume the same rectangular shape, $\Delta_x \times \Delta_y$. In practice, square pixels are common, $\Delta_x = \Delta_y = \Delta$. Each pixel is uniquely identified by its Cartesian location. We may assume, specifically, that the center of the (i, j) pixel is at $(i\Delta_x, j\Delta_y)$, $-M \leq i \leq M$, $-N \leq j \leq N$, with $(2M + 1)\Delta_x = X$, $(2N + 1)\Delta_y = Y$. The characteristic function of the (i, j) pixel at the detector array is

$$P_{ij}(x, y) = \text{rect}\left(\frac{x}{\Delta_x} - i\right) \text{rect}\left(\frac{y}{\Delta_y} - j\right), \quad (2)$$

where $\text{rect}(x) = 1$ if $x \in [-1/2, 1/2]$ and $\text{rect}(x) = 0$ otherwise. The pixel function, P_{ij} , described above represents a unity fill factor pixel. In practice, this ideal function has to be revised to describe incomplete fill factors in actual electronic pixels. When there is no additional coding at the focal plane, the pixel sampling function can be simply described by multiplication of the pixel function and the function to sample from

$$s_{ij}(x, y, \tilde{f}(x, y)) = P_{ij}(x, y) \tilde{f}(x, y).$$

With such a sampling function, the pixel at (i, j) location is said to be *clear*. We shall introduce a couple of cases where nontrivial focal-plane coding schemes are employed together with the use of multiapertures.

In the general case of multiapertures, the model in Eqs. (1) applies to each and every individual aperture.

B. Lenslet Displacements

The TOMBO system [9] can be characterized as a special case of the system model (1). It aims at reducing the thickness of the optical system by replacing a single large aperture lens with an array of smaller aperture lenses or lenslets. The detector pixel array is accordingly partitioned so that each and every lenslet corresponds to a subarray. We refer to the image on a subarray as a subimage, relative to that on the entire array. In order to maintain the resolution of the system with a single large aperture lens, a diversity in the subimages is essential. Otherwise, the detector array carries redundant, identical subimages of lower resolution.

The TOMBO system exploits the relative nonuniform displacements of the lenslets at manufacture. Consider a 3×3 array of lenslets. There are nine subimages at 3×3 subapertures, $G_{pq} = [g_{pq,ij}]$, and $p, q = -1, 0, 1$. Here, the double indices pq associated with G in the capital case specify the subarray or subimage corresponding to the lenslet at location p, q ; the four-tuple indices (pq, ij) associated with g in lowercase specify the (i, j) pixel in the (p, q) subarray or subimage. The subimage at the center is G_{00} . The subimage G_{pq} is made different from G_{00} by the displacement of the (p, q) lenslet relative to the center one.

By the COMP-I concept, the lenslet displacement can be *designed* first and calibrated after manufacture. In terms of the framework in Subsection 2.A, we have the following specific model:

$$g_{pq,ij} = \int_{-X/2}^{X/2} \int_{-Y/2}^{Y/2} s_{pq,ij}(x, y, \tilde{f}(x, y)) dx dy, \quad (3)$$

$$\tilde{f}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) h(\beta x - \xi, \beta y - \eta) d\xi d\eta,$$

where β is the scaling factor relating the object scene to the subapertures. The choice $\beta = 3\alpha$ is natural for

the 3×3 lenslet array, where α is the scaling factor for a comparison single entire aperture system. The sampling functions for lenslet displacements can be described as follows:

$$s_{pq,ij}(x, y, \tilde{f}(x, y)) = P_{ij}(x, y)E_{pq}(\tilde{f}(x, y)) \\ = P_{ij}(x, y)\tilde{f}(x - \delta_p, y - \delta_q), \quad (4)$$

where E denotes the shift operator and δ_p and δ_q are not multiples of Δ . By design, one may deliberately set $\delta_p = p\delta$, $\delta_q = q\delta$ with $\delta = \Delta/3$.

A couple of remarks are in order. It is assumed that the lenslets have the identical views of the same object at the scene. The lenslet shifting, which is non-circulant, requires the additional knowledge or assumption of the boundary conditions in numerical reconstruction. We will discuss the procedures for computational integration of the multiple images at the multiple apertures and their numerical properties. We show next that the sampling functions in the array model (3) are not restricted to the lenslet displacement in theory and practice.

C. Codified Focal-Plane Masks

The model framework of Eqs. (1) and (3) admits of various focal-plane coding schemes. We introduce another focal-plane coding scheme, which is implemented in COMP-I as well. Each lenslet is associated with a mask representing a unique code. Consider, for example, a 4×4 lenslet array. Figure 1 shows all 16 masks, one for each lenslet, based on what we refer to as the modified Hadamard code. Every subaperture has a unique mask; each and every pixel in the subaperture is masked with the same pattern. The pixelwise sampling function can be described as follows:

$$s_{pq,ij}(pq, ij)(x, y, \tilde{f}(x, y)) = P_{ij}(x, y)\mathcal{H}_{pq}(x, y)\tilde{f}(x, y), \quad (5)$$

where the function \mathcal{H}_{pq} represents the codeword im-

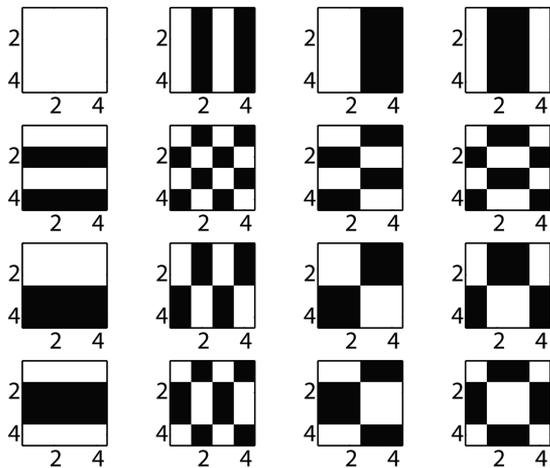


Fig. 1. Focal-plane pixel masks for a 4×4 lenslet array based on the modified Hadamard coding scheme.

plemented by the mask at the (p, q) lenslet or subaperture. The first subaperture is clear or unmasked.

The masks, or two-dimensional codewords, are drawn from Hadamard matrices. The masks on a 4×4 lenslet array are particularly constructed as follows:

$$\mathcal{H}_{pq} = \frac{1}{2}((H_4 e_q)(H_4 e_p^T) + ee^T), \quad (6)$$

where H_4 is the 4×4 Hadamard matrix with elements in $\{-1, 1\}$ (see the Appendix), e_k is the k th column of the identity matrix, and e is the column vector with all elements equal to 1. In other words, we take the q th column and p th row of H_4 , form the outer product, and replace -1 with 0 by lifting the elements by 1 and scaling them by $1/2$. This results in the masks shown in Fig. 1, where the black subpixels block the light and represent zero values in the codewords \mathcal{H}_{pq} . Masks with binary patterns are relatively easy to manufacture, which we detail in Section 3. The masks have the effect of partitioning every single pixel into 4×4 subpixels.

The order of Hadamard matrices is limited to powers of 2. We lift this limitation by introducing 0–1 valued and well-conditioned matrices such as

$$H_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \quad H_5 = \begin{pmatrix} H_4 & e \\ e^T & -1 \end{pmatrix}. \quad (7)$$

They are in fact the leading submatrices of H_4 and H_5 , respectively.

The sampling function specified pixelwise by Eqs. (5) and (6) is well conditioned for numerical image reconstruction. As we shall show shortly, the modified Hadamard focal-plane coding scheme is much more efficient in computational reconstruction and much less subject to numerical instability.

D. Decoding Analysis

We analyze the decoding or integrating process, which is an important step for computational construction of a single image at the subpixel level, compensating the coarse resolution with the thin optics. It is worth noting that the decoding process precedes the conventional PSF deconvolution.

Assume a $P \times Q$ lenslet array. The pixel array A is partitioned into $P \times Q$ blocks, each block corresponding to a subimage with $M \times N$ pixels. We use $A(p, q, i, j)$ to specify the (i, j) element in the (p, q) subarray. We introduce a related array \hat{A} ,

$$\hat{A}(i, j, p, q) = A(p, q, i, j). \quad (8)$$

In aggregated array form, $\hat{A} = P_{MP,M} A P_{NQ,N}^T$, where $P_{n,k}$ denotes a stride- k permutation of n elements. The permuted array \hat{A} is of $M \times N$ blocks, each block with $P \times Q$ pixels. In detail, the (i, j) block in \hat{A} is composed of the (i, j) pixels from all the subarrays in A . In

other words, by the stride permutation, we integrate the subimages into a single image at the level of geometric pixel blocks, and the pixels in each geometric block correspond to the same sample portion of the image f . We may say simply that the block array \hat{A} is a multiplex of the subimages, and the block array A is the image with multiplex pixels. If all the pixels in the (i,j) block of \hat{A} were measured without being shifted or masked, they would have assumed identical values in the absence of any distortion. The single image would be one at the coarse level with the $P \times Q$ sensor pixel constituting an image pixel. Focal-plane coding is an approach for restoring the resolution.

The process of decoding for a system with the modified Hadamard codes is highly computationally efficient. The integration of multiple subimages at a coarse resolution level to a single image of higher resolution is direct and local within each geometric pixel block in the permuted array \hat{A} , after the global permutations.

Theorem 1 Assume a $P \times Q$ lenslet array with the modified Hadamard coding (6). Then, the sampling function $P_{ij} \mathcal{H}_{pq}$ partitions the (i,j) pixel into $P \times Q$ subpixels. Denote by X_{ij} the pixelwise image $P_{ij}f$. Let M_{ij} be the corresponding $P \times Q$ pixel block in A , $M_{ij} = \hat{A}(i,j, :, :)$, where \hat{A} is defined in Eq. (8). Then,

$$2M_{ij} = H_p X_{ij} H_q + (e^T X_{ij} e) e e^T, \quad (9)$$

$$X_{ij} = H_p^{-1} (2M_{ij} - M_{ij}(1,1) e e^T) H_q^{-1}. \quad (10)$$

The block array $X = [X_{ij}]$ renders the single integrated image.

In the theorem, Eq. (9) describes the coded measurements of each pixelwise image, and Eq. (10) is for decoding per pixel block. In the inversion process we use the fact that the pixel $M_{ij}(1,1)$ is clear and therefore $M_{ij}(1,1) = e^T X_{ij} e$. We omit the detailed proof.

In addition to simplicity and efficiency, the decoding process is well conditioned. The condition number of a matrix is the ratio of its largest to smallest eigenvalues in absolute value. It is a metric that can be used to quantify the worst-case loss of precision with the presence of noise in measured data and truncation errors in image construction. Smaller condition numbers indicate less sensitivity to noise or error in the image reconstruction. Large condition numbers on the other hand suggest the potential of severe degradation in image resolution. The condition number for the image reconstruction from a $p \times q/2$ lenslet array is amazingly small; it is approximately PQ , a half of the number of lenslets.

Corollary 2 Assume the modified Hadamard focal-plane coding for a $P \times Q$ array with P and Q as powers of 2. Then the condition number for the modified Hadamard coding and single-image construction is

$$\text{cond}(S_{PQ}) = \frac{(PQ + 4) + \sqrt{(PQ + 4)^2 - 4^2}}{4}. \quad (11)$$

We leave the proof to the Appendix. In Table 1 we provide the condition numbers for $K \times K$ arrays with K from 2 to 8. Whenever K is not a power of 2, matrix H_K is the $K \times K$ leading submatrix of H_8 . These condition numbers are all modest, depending only on the partition of the sensor array.

In comparison, the coding by lenslet displacements entails quite a different kind of decoding process. For simplicity in notation, the statements in the following theorem are for a lenslet array with equally spaced subpixel displacements.

Theorem 3 Assume a $P \times Q$ lenslet array with lenslet displacements in step size $\delta_x = \Delta_x/P$ in x dimension and $\delta_y = \Delta_y/Q$ in y dimension. Assume the sensor array is composed of $M \times N$ pixels. Let \hat{A} be the entire sensor array permuted as in Eq. (8). Each pixel in the sensor array is partitioned by the sampling function into $P \times Q$ subpixels. Denote by X the entire image at the subpixel level.

Part I. The mapping between X and \hat{A} is globally coupled

$$\hat{A} = B_{P,M} X B_{Q,N}, \quad (12)$$

where $B_{P,M}$ is an $M \times M$ Toeplitz 0-1-valued matrix with P diagonals when the boundary values are zero.

Part II. The condition number for the decoding process depends on not only the lenslet partition $P \times Q$ but also the size of the entire detector array $M \times N$ and the boundary condition.

Part I of Theorem 3 describes the global coupling relationship between \hat{A} and X induced by the lenslet displacements, in contrast to the pixelwise coupling by the modified Hadamard coding. The bands of the coupling matrices in both dimensions may shift to the right or left, depending on the assumed location of the center subaperture. In addition, it assumes zero values at the shifted-in subpixels. The following corollary gives a special case of the statements in part-II.

Corollary 4 Assume the conditions of Theorem 3. Assume in addition that $P = Q = 3$, $M = N$, which is a multiple of 3, the lenslet array is centered at the middle lenslet, and the boundary values are zero. Then, $B_{3,N}$ is symmetric and tridiagonal. The condi-

Table 1. Condition Numbers for the Decoding Process Associated with the Modified Hadamard Coding on $K \times K$ Lenslet Arrays

K	2	3	4	5	6	7	8
Condition #	3.7	11.8	9.9	28.8	24.8	50.8	34.0

tion number for decoding is bounded as follows:

$$\left(\frac{1 + 2 \cos\left(\frac{\pi}{N+1}\right)}{1 + 2 \cos\left(\frac{[2(N+1)/3]\pi}{N+1}\right)} \right)^2 < \text{cond}(B_{3,N} \otimes B_{3,N}) < \left(\frac{1 + 2 \cos\left(\frac{\pi}{N+1}\right)}{1 + 2 \cos\left(\frac{[2(N+1)/3]\pi}{N+1}\right)} \right)^2. \quad (13)$$

We leave a proof sketch in the Appendix. We give in Table 2 the conditions numbers for a few cases, which increase with the number of pixels in a subimage.

The substantial difference in the sensitivity to noise in decoding between the two coding schemes, as shown in Tables 1 and 2, has a significant implication. The decoding process for the lenslet-displacement scheme with a large sensor array may have to resort to iterative methods, as in TOMBO [9,15], because the computation by direct solvers takes more operations and memory space, with few exceptional cases. Thus, one needs to determine effective strategies for initial values, iteration steps, and termination. In this type of shift-coded system, significantly large numerical errors may result from a decoding procedure, thus introducing an additional source of errors. Moreover, the decoding process is also sensitive to any change in the boundary condition. These problems do not exist in the image integration process with the modified Hadamard scheme, a direct method based on the explicit expression in Eq. (10).

The coding analysis demonstrated in this section underlines the design of a focal-plane coding system from the aspect of computational image integration. In particular, it has added more to our understanding of the TOMBO system. The decoding process is only a step of the single-image reconstruction process, in addition to the conventional steps. We will describe the reconstruction in more detail in Section 5.

3. System Implementation

There is a distance from the theoretical coding design to its implementation. We describe in this section the technical challenges and our resolutions. Briefly, a stock camera board was modified with a customized focal-plane coding element. For the optics, a custom lenslet array was manufactured by Digital Optics Corporation, a subsidiary of Tessera Technologies.

A. Focal-Plane Array

A Lumenera Lu100 monochrome board level camera is used for data acquisition. Built on an Omnivision OV9121 imaging sensor, the focal-plane array con-

sists of 1280×1024 pixels, with each pixel of $5.2 \mu\text{m} \times 5.2 \mu\text{m}$ in physical size. The camera uses complementary metal-oxide semiconductor technology where each pixel contains a photodiode and an individual charge to voltage circuitry. These additional electronics reduce the light-sensitive area of a pixel. However, each pixel has a microlens that improves photon collection. These microlenses provide a nonuniform improvement as a function of field angle of the incident light. We recognize that our current model assumes a response independent of the light's incident angle on the detector.

The conventional imaging sensor is isolated from the environment with a thin piece of glass by the manufacturer. However, we need to place the focal-plane coding element in direct contact with the imaging sensor. It was challenging to remove the glass without damaging the pixels underneath. We developed a procedure to dissolve the adhesive securing the cover glass. We applied an acetone and ethyl ether mixture around the perimeter of the sensor. At the same time, we scraped at and removed the residual of the adhesive with a razor blade edge. Complete removal of the cover glass required multiple chemical applications.

B. Lenslet Array

The lenslet array used in the COMP-I imager is a hybrid of two refractive surfaces and one diffractive surface per lenslet. The refractive lenses are fabricated using lithographic means on two separate 150 mm wafers made of fused silica. The final lens shapes are aspheric. On the wafer surface opposite one of the lenses, an eight-phase level diffractive element is fabricated using the binary optics process. The diffractive element primarily performs chromatic aberration correction. The two wafers, one refractive and the other with a refractive and diffractive surface, are bonded together, with the two refractive surfaces facing away from each other, via an epoxy bonding process. A spin-coated and patterned spacer layer of $20 \mu\text{m}$ thickness controls the gap between the wafers. After bonding, a dicing process singulated the wafer. Figure 2 shows the unmounted lenslet array.

When integrated, the distance from the front lens surface to the focal plane is approximately 2.2 mm. The imaging system functions as an $f/2.1$ lens with an EFL of 1.5 mm. Centered at 550 nm, the system operates principally over the green portion of the visible spectrum. A patterned chrome coating on the first refractive surface of the optic is the limiting aperture. Prior to chrome deposition, a dielectric coating placed on the first surface acts as an IR cut filter.

C. Focal-Plane Coding Element

The focal-plane coding element is a patterned chrome layer on a thin glass substrate fabricated with reduction lithography by Applied Image, Inc. The registration of the focal-plane coding element

Table 2. Condition Numbers Associated With the $P \times P$ Partition of $N \times N$ Detector Arrays

P	N	Condition Number
3	3	$4.86e + 05$
5	5	$4.04e + 05$
1280	1024	$1.26e + 06$

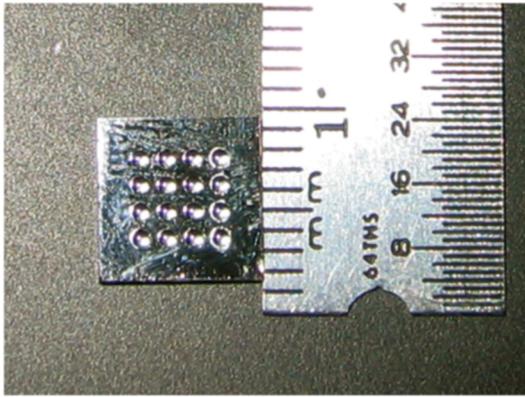


Fig. 2. Unmounted refractive lenslet array.

with the pixel axis is important. Proper alignment with the pixel axis used the nonimaging perimeter pixels of the 1280×1024 sensor. Specifically, we designed and patterned subpixel sized placement marks on the border outside the central 1000×1000 pixels on the glass substrate. The feature size on the mask is $1.3 \mu\text{m}$, designed to be one quarter of the camera pixel. Figure 3 shows these marks under magnification. Figure 4 shows two coding regions of the focal-plane element.

We developed a process to affix the glass substrate to the imaging sensor. A vacuum helped to hold the mask stationary, and we positioned the camera board directly under it with 100 nm accuracy Newport AutoAlign positioning equipment. First, we aligned the mask to completely reside within the active area of the imaging sensor. Next, the stages aided to decrease the gap between the glass and the detector, and the vacuum was then turned off. A small needle dispensed a drop of UV-curable adhesive on the vertical edges of the glass. In sufficient time, capillary action drew a very thin layer of the viscous adhesive between the glass substrate and the imaging sensor. In an active alignment process, captured images

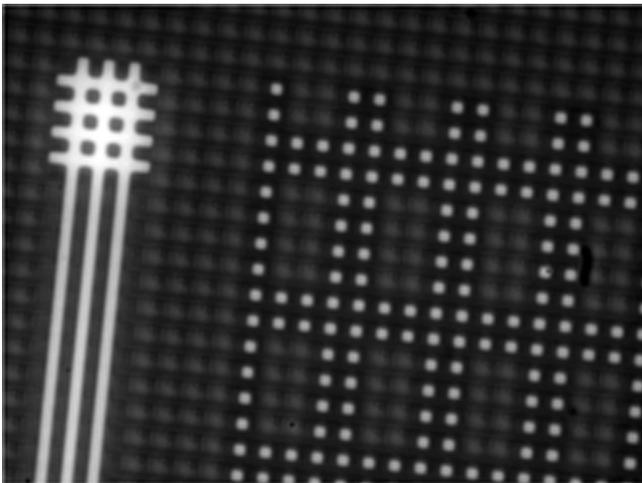


Fig. 3. Microscope image of the imaging sensor with a focal-plane coding element. The white dots and bars are alignment marks on the focal-plane coding element.

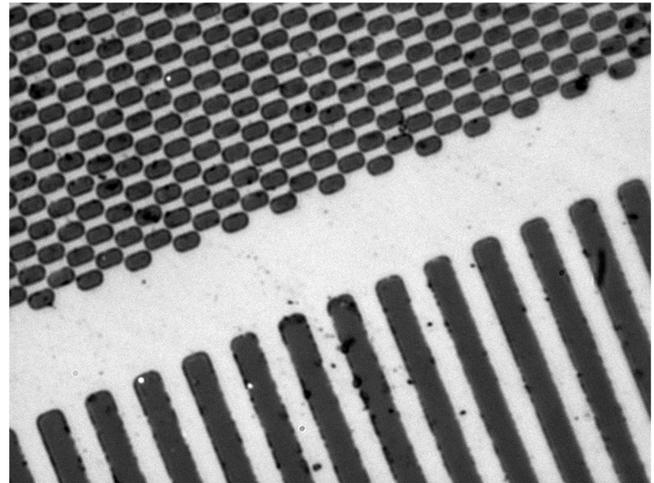


Fig. 4. Focal-plane coding element under $100\times$ magnification. Two aperture patterns are visible. The period of the bottom grating is equal to the pixel pitch.

guided the registration the mask features with the pixels. The tip of the adhesive distribution needle nudged the mask to its final position. Finally, a high intensity UV lamp cured the adhesive to secure the mask.

D. Lens Alignment

Alignment of the lenslet element with the focal plane is a major challenge in the system integration. With a focal length on the millimeter scale, the depth of focus for these lenses is of the order of micrometers. This requires very precise translation methods and very narrow tolerances. Additionally, a second problem exists in that determining the system's best focus is not trivial.

In order to hold the optics, we specified the dimensions of a custom lens mount with computer-aided design software. An Objet Eden 330 rapid prototyping machine printed the part using a stereolithography technique. A 6-axis precision positioning system adjusted the camera board with respect to the stationary lens.

In order to align the focus, the lenslet array images a bright point source onto the detector. The source is placed on axis and in the far field at a distance of well over 100 focal lengths. One traditionally determines that a system is in focus when a point source in the object plane produces the most concise PSF. In our system, the spatial extent of the PSF is smaller than a pixel. Our lens design very nearly reaches the diffraction limit, with PSF radius equal to $1.22\lambda f/d$, where λ is the wavelength, f is the focal length, and d is the lens diameter.

Determining the best focus for a system is challenging when the desired spot size is smaller than a pixel. The PSF width cannot be easily measured because the pixels electronically downsample the optical field. In order to attack this problem, we captured sequential images while translating the camera along the optical axis. Qualitatively, when the image is out

of focus, one observes light intensity spread over multiple pixels, and as the focal spot becomes smaller, the intensity becomes more localized to a single pixel. Numerically, one metric we employed is the calculation of the standard deviation of pixel value intensities in a cropped region surrounding the spot. When the spot is out of focus, one expects to see a lower standard deviation because of the more uniform distribution. If the spot is in focus, nearly all intensity is on a single pixel and the calculated standard deviation is much higher.

A potential complication is the possibility that the system is aligned in such a way that, when in best focus, an impulse falls centered on the border between 2 (or 4) pixels. The resulting captured image would still show intensity split between those pixels, even though the spot size is smaller than a single pixel.

However, the more interesting problem is determining the best focus for apertures with a focal-plane coding element. If a point source images to a masked region of the detector, one would expect to see minimal response when the system is in best focus. Furthermore, if the spot size grows, it could potentially increase the pixel response of a given camera, with minimal effect on neighboring pixels. Thus, the result would appear nearly identical to a situation where the system is in best focus imaging a point source to an unmasked region on the detector. In order to differentiate between the two, one needs to translate the image with respect to the camera pixels.

4. Impulse Response Measurement

The focal-plane coding element modulates pixel responses differently in each aperture. Since the period of the mask pattern is equal to the pixel spacing, pixels in a given aperture all share identical modulation characteristics. However, determining the exact registration of the mask with the camera pixels requires calibration.

A point source is approximated by a fiber coupled white light source placed in the far field. When translating the focal-plane array perpendicular to the optical axis, the image of the point source moves correspondingly. We capture images at multiple point source locations. A computerized translation achieves subpixel positioning of the point source on the detector. In Eq. (1), $f(x, y) \approx \delta(x, y)$ represents a point source. Thus, we essentially measure the convolution of the lens's PSF with the sampling function of the detector.

We will first examine the aperture with a binary grating with a 50% duty cycle. The pattern is uniform in the horizontal direction. Facing the camera, this code appears in the lower left lenslet. Figure 1 shows the designed focal-plane code on each pixel. For the scan, we translated the center of the point source vertically in increments of $0.2 \mu\text{m}$ compared with the $5.2 \mu\text{m}$ pixel pitch. Figure 5 shows four adjacent pixels'

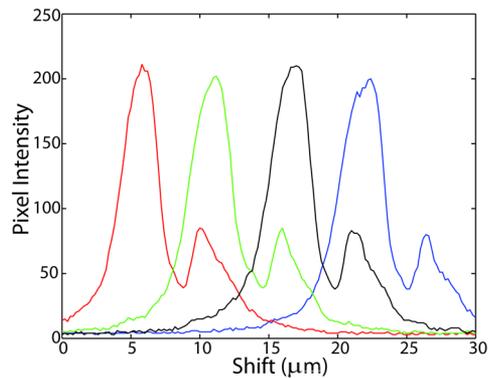


Fig. 5. Impulse response scan of four adjacent pixels. Each line corresponds to a pixel's intensity as a function of the relative position of a point source on the detector plane.

els' impulse responses as a point source is translated. Each line plots the response of a pixel as a function of the relative movement of the point source on the detector. The response gradually shifts from one pixel to the next as the center of the point source tracks from one pixel to its neighbor. The width of the impulse response measurement is broader than the $5.2 \mu\text{m}$ pixel pitch because of the finite extent of the PSF.

We took this same impulse response measurement for a two-dimensional array of point source locations. Here, translation stages positioned a point source perpendicular to the optical axis in a two-dimensional grid as images are captured from the camera. A typical scan might consist of 100×100 object locations covering approximately a 5×5 pixel square in image space.

Figure 6 shows impulse response data captured from an unmasked pixel. The asymmetric nature of the complementary metal-oxide semiconductor

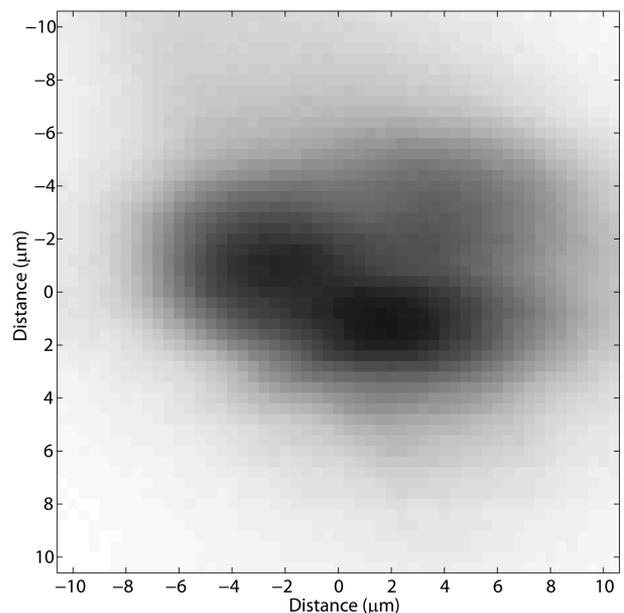


Fig. 6. Pixel impulse response as a function of image location on detector plane.

pixel's sampling is most likely a result of the pixel's lack of sensitivity where the charge to voltage circuitry resides. As expected, there is only minimal variation between responses across apertures. This was verified by observing nearly identical responses for pixels *within a single subaperture*. While Fig. 6 shows just a single pixel, data were collected for its neighbors and inspected visually for consistency. The impulse response is shift invariant on a macro scale from pixel to pixel but shift variant within a single pixel.

Even more interesting, though, is the modulation of the impulse response shown in Figs. 7 and 8. The focal-plane coding element's effect is clearly visible. The pixel exhibits a modified response due to the subpixel mask features. It is important to note again that a precondition of this result is that the PSF of the optical system is smaller than the features on the coding mask. Without such a well-confined spot, the mask would not have such a significant effect. A larger spot would imply a narrower extent in the Fourier domain and would essentially low-pass filter the aperture sampling function. An impulse response shape similar to the open aperture would be observed because the mask features (at higher spatial frequencies) would be attenuated.

5. Single Image Construction

We describe the computational construction of a single image from multiple images on the sensor subarrays. In addition to the conventional steps of noise removal and deblurring, the reconstruction has the distinct decoding step for integrating multiple images into a single image. We illustrate the

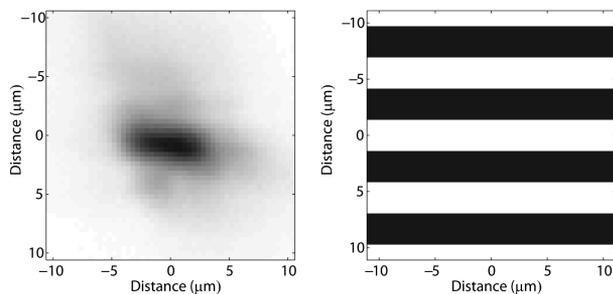


Fig. 7. Impulse response of a pixel masked with a 50% horizontal grating with period equal to the pixel pitch.

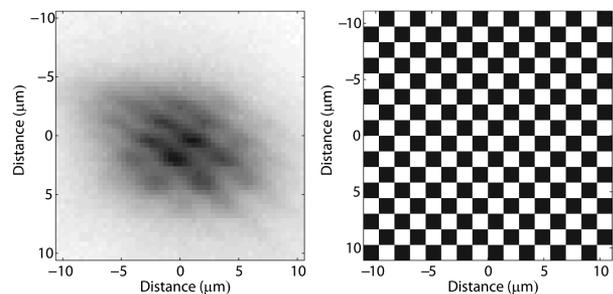


Fig. 8. Impulse response of a pixel masked with a checkerboard with feature size equal to one quarter pixel.

reconstruction procedure with the particular case of the modified Hadamard coding scheme.

By the model (1), the subimage registered at each subaperture is considered as a projection of the same optical information along a different sampling channel. The construction of a single image from the multiple projections is therefore a backprojection for image integration. The single-image construction consists of three major stages. The first two stages prepare for the final backprojection stage or the decoding stage.

First, every subimage corresponding to a lenslet is individually cropped from the mosaic image captured at the sensor array. The subimage is registered according to the sensor pixels associated with the lenslet. This cropping step requires the calibration of the subarray partition and alignment. The calibration may be done once for all in an ideal case or carried out periodically or adaptively otherwise. Figure 9 shows the raw mosaic image, captured by the camera, of the ISO-12233 Digital Still-Camera Resolution Chart.

Second, each and every subimage is filtered individually for noise removal, as in conventional image processing. The additional task in this stage is the adjustment of the relative intensity (in numerical values) between the subimages.

The final decoding stage follows Theorem 1. In terms of procedural steps, the subimages are first integrated pixel block by pixel block. Specifically, for a $P \times P$ lenslet array, the (i, j) pixel block is composed of the (i, j) pixels from $P \times P$ subimages. Next, local to each and every pixel block, a block image of $P \times P$ subpixels is constructed by using the explicit formula (10). Both steps are local to each pixel block, or parallel among pixels blocks. This simple, deterministic, and parallel process has great potential to be easily embedded into imaging systems using, for example, field programmable gate array hardware. We omit a



Fig. 9. Raw captured image from the multiple-aperture focal-plane-coded camera. Here, the target is a portion of an ISO-12233 Digital Still-Camera Resolution Chart.

detailed discussion of such embedding because it is beyond the scope of this paper.

Figure 10 shows details of a reconstructed image (on the left) compared with the bicubic spline interpolation of the clear aperture lenslet subimage (on the right) resized from 128×128 pixel cropped image to 512×512 pixel size of the reconstruction.

6. Conclusion

The success of a high-resolution optical imaging system equipped with a focal-plane coding scheme relies on the integrated design and analysis of the coding and decoding schemes with full respect to the potential and limitation of the physical implementation and numerical computation. We have presented a framework of focal-plane coding schemes for multichannel sampling in optical systems. Focal-plane coding is a sampling strategy that modulates the intrinsic pixel response function. Among other feasible schemes in the framework, we discussed lenslet displacements and coding masks. In the former scheme, the displacement pattern can be determined by primary design and further calibration. The latter scheme has advantages in computational efficiency and stability. While masks block photons, one could avoid this loss by designing more complex sampling functions in the focal plane. We implemented both coding schemes in the COMP-I project.

Conventional systems image the scene onto the detector and sample that distribution with a pixel array. These systems typically use the raw pixel intensity as the estimate for the image in that location, resulting in a sampling rate directly related to pixel pitch. The COMP-I system does not use this sampling approach. By integrating the multichannel data we can achieve a smaller effective pixel size than what is measured in the individual subimages. In our system we are not attempting to deconvolve the optical PSF, and we recognize that the best image we can recover is the one that reaches the detector before any electronic sampling. The difference is that we are attempting to virtually subdivide raw electronic pixels by applying a different coding pattern (or sampling function) to each subaperture (or channel). This is possible because the coding mask has features on the scale of this virtual pixel size.

We have detailed a thin camera with a lenslet array and focal-plane coding element to mask each

lenslet differently. We also designed custom optics for the imaging system and developed their alignment procedures. The system was tested, and we showed that the coding masks have the designed functionality.

Appendix

The Kronecker product. Let A be an m -by- n matrix and let B be a p -by- q matrix. Then the Kronecker product $A \otimes B$ is defined as the mp -by- nq block matrix:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}. \quad (\text{A1})$$

Hadamard matrices. The Hadamard matrices are conventionally defined at powers of 2 in the matrix order by the recursion

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad H_{2n} = H_2 \otimes H_n, \quad (\text{A2})$$

$$n = 2^k, \quad k \geq 1.$$

Proof of Corollary 2. First, the system of linear equations in Eq. (9) can be expressed equivalently in a one-dimensional format as follows:

$$\text{vec}[M_{ij}] = \frac{1}{2}[(H_P \otimes H_Q) + ee^T] \text{vec}[X_{ij}], \quad (\text{A3})$$

where $\text{vec}[\cdot]$ sequentializes a two-dimensional array column by column into a one-dimensional vector. When P and Q are powers of 2, $H_P \times H_Q$ is a conventional Hadamard matrix. It remains to determine the singular values of the modified Hadamard matrix $\tilde{H}_N = (H_N + ee^T)/2$, where N is a power of 2. We have

$$\tilde{H}_N = \frac{1}{2}H_N \left(I + \frac{1}{N}H_N^T ee^T \right) = \frac{1}{2}H_N (I + e_1 e_1^T).$$

The condition number of \tilde{H}_N is the same as that of $(I + e_1 e_1^T)$, or the square root of the condition number of $(I + e_1 e_1^T)(I + ee_1^T)$. The symmetric and positive definite matrix is similar by orthogonal transformation to the following matrix:

$$\begin{pmatrix} N+3 & \sqrt{N-1} & 0 \\ \sqrt{N-1} & 1 & 0 \\ 0 & 0 & I_{N-2} \end{pmatrix}.$$

It is then straightforward to verify that the two extreme eigenvalues are

$$\lambda_{\max} = \frac{N+4 + \sqrt{(N+4)^2 - 4^2}}{2} > 1 > \lambda_{\min}$$

$$= \frac{N+4 - \sqrt{(N+4)^2 - 4^2}}{2} > 0.$$

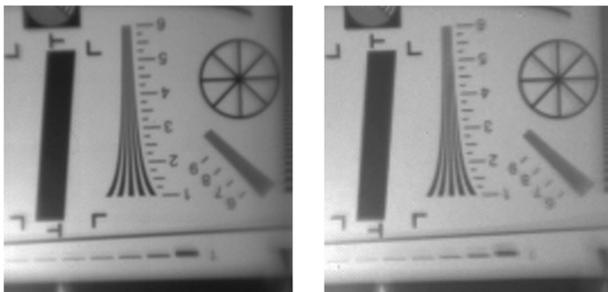


Fig. 10. Detail of the reconstruction (left) compared with the bicubic spline interpolation of the clear aperture image (right).

Taking the square root of $\lambda_{\max}/\lambda_{\min}$, we have the condition number of \tilde{H}_N as given in Corollary 2.

Proof of Corollary 4. The eigenvalues of the $N \times N$ tridiagonal matrix with -2 on the diagonal and 1 on the superdiagonal and subdiagonal are known analytically as $-4\sin^2(j\pi/2(N+1))$, $j = 1, 2, \dots, N$; see [16] for example. The eigenvalues shifted by 3 are those of the tridiagonal matrix $B_{3,N}$.

This work is supported in part by Microsystems Technology Office of Defense Advanced Research Projects Agency contract N01-AA-23103 and by the National Science Foundation. The authors thank all the COMP-I team members at Duke University, the University of North Carolina at Charlotte, the University of Delaware, the University of Alabama in Huntsville, the Army Research Laboratory, and Digital Optics Corporation, a subsidiary of Tesser Technologies, Inc.

References

1. A. Papoulis, "Generalized sampling expansion," *IEEE Trans. Circuit Syst.* **24**, 652–654 (1977).
2. N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Trans. Image Process.* **10**, 573–583 (2001).
3. S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.* **20**, 21–36 (2003).
4. H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP: Graph. Models Image Process.* **54**, 181–186 (1992).
5. M. A. Fiddy and T. J. Hall, "Nonuniqueness of superresolution techniques applied to sampled data," *J. Opt. Soc. Am.* **71**, 1406–1407 (1981).
6. R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.* **37**, 247–260 (1998).
7. Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 83–97 (2004).
8. Y. Lu and M. Inamura, "Super-resolution of undersampled and subpixel shifted image sequence by pyramid iterative back-projection," *IEICE Trans. Inf. Syst.* **E85D**, 1929–1937 (2002).
9. J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka, "Thin observation module by bound optics (TOMBO): Concept and experimental verification," *Appl. Opt.* **40**, 1806–1813 (2001).
10. D. J. Brady, M. Feldman, N. Pitsianis, J. P. Guo, A. Portnoy, "Compressive optical montage photography," *Proc. SPIE* **5907**, 590708 (2005).
11. N. P. Pitsianis, D. J. Brady, and X. Sun, "Sensor-layer image compression based on the quantized cosine transform," *Proc. SPIE* **5817**, 250 (2005).
12. A. D. Portnoy, N. P. Pitsianis, D. J. Brady, J. Guo, M. A. Fiddy, M. R. Feldman, and R. D. T. Kolste, "Thin digital imaging systems using focal-plane coding," *Proc. SPIE* **6065**, 60650F (2006).
13. R. Mait, J. Athale, and J. van der Gracht, "Evolutionary paths in imaging and recent trends," *Opt. Express* **11**, 2093–2101 (2003).
14. N. P. Pitsianis, D. J. Brady, A. Portnoy, X. Sun, M. A. Fiddy, T. Suleski, M. R. Feldman, and R. D. TeKolste, "Compressive imaging sensors," *Proc. SPIE* **6232**, 43–51 (2006).
15. J. Tanida, R. Shogenji, Y. Kitamura, K. Yamada, M. Miyamoto, and S. Miyatake, "Color imaging with an integrated compound imaging system," *Opt. Express* **11**, 2109–2117 (2003).
16. C. Van Loan, *Computational Frameworks for the Fast Fourier Transform* (Society for Industrial and Applied Mathematics, 1992).