# Compressive Optical MONTAGE Photography

David J. Brady[a], Michael Feldman[b], Nikos Pitsianis[a], J. P. Guo[a], Andrew Portnoy[a], Michael Fiddy[c]

[a]Fitzpatrick Center, Box 90291, Pratt School of Engineering, Duke University, Durham, NC 27708
[b]Digital Optics Corporation, 9815 David Taylor Drive, Charlotte, NC 28262
[c]Center for Optoelectronics and Optical Communications, University of North Carolina Charlotte, 9201 University City Blvd. Charlotte, NC 28223.

## ABSTRACT

The Compressive Optical MONTAGE Photography Initiative (COMP-I) is an initiative under DARPA's MONTAGE program. The goals of COMP-I are to produce 1 mm thick visible imaging systems and 5 mm thick IR systems without compromising pixel-limited resolution. Innovations of COMP-I include **focal-plane coding**, **block-wise focal plane codes**, **birefringent, holographic and 3D optical elements for focal plane remapping** and **embedded algorithms for image formation**. In addition to meeting MONTAGE specifications for sensor thickness, focal plane coding enables a reduction in the transverse aperture size, physical layer compression of multispectral and hyperspectral data cubes, joint optical and electronic optimization for 3D sensing, tracking, feature-specific imaging and conformal array deployment.

**Keywords:** Focal plane, multiple aperture, transmission masks, image sampling

## 1. INTRODUCTION

Focal plane coding, consisting of structured arrangement of pixel sampling geometry, enables non-degenerate sampling in multiple aperture imaging systems. The COMP-I program uses of transmission masks for focal plane coding and has shown that focal plane coding enables compressive image sampling. Core aspects of COMP-I systems include

1. ***Focal plane coding*** is the core COMP-I innovation. The fundamental vision of MONTAGE is to use integrated sensing and processing to break the conventional isomorphism between image pixels and digital samples. Focal plane coding consists of intelligent sampling and remapping of optical pixels to enable efficient digital reconstruction. COMP-I has explored four approaches to focal plane coding: focal plane pixel masks, birefringent and refractive remapping elements, holographic lenslet arrays and photonic crystal remapping elements. In first generation systems, focal plane coding is the preferred implementation strategy.

2. ***Block-wise and multiscale focal plane codes*** have been developed by the COMP-I team. Block-wise and wavelet coding form the basis of current image compression schemes. The COMP-I team has developed and simulated codes based on realistic optical design rules that improve the SNR for image estimation from generalized sampling by orders of magnitude and that enable data efficient transverse sampling.

3. ***Compressive imaging*** is an extension of generalized sampling to include the sampling on image-aware bases. Compressive sampling measures non-local bases using focal plane coding to enable sampling below naïve Nyquist limits. The COMP-I team has demonstrated the feasibility of compressive imaging in simulation of focal plane coded systems.

4. ***Image fusion, inversion and registration algorithms*** are enabled by block-wise and multi-resolution approaches. Because individual sub-apertures or clusters of apertures may be designed to reconstruct the full resolution optical image, image fusion for higher level estimation may rely on Bayesian or other nonlinear estimation algorithms on the full resolution image. This approach dramatically relaxes registration criteria.

5. ***"Thin imaging"*** consistent with the MONTAGE specification is enabled by innovations 1-5. COMP-I achieves 5-10 times reduction in system thickness through focal plane codes to enable sub-Nyquist transverse sampling of the optical intensity distribution on the focal plane. COMP-I achieves 6-9 times reduction by joint optimization of physical design.

Imaging system design begins with the focal plane. Suppose that the focal plane consists of pixels of size $\delta$. The size of the image on the focal plane is D. The number of pixels is $N = D/\delta$. In a conventional imaging system, a lens of focal length $F = D/N.A.$ is used to form an image. The diffraction limited resolution of the field distribution on the focal plane is $\lambda/N.A.$. Typically, the diffraction limited resolution is much less than $\delta$. The angular field of view is approximately $\sin\theta = D/F = N.A.$. The angular resolution is $\Delta\theta_\delta = \delta/F$ due to the focal plane and $\Delta\theta_\lambda = \lambda/D$ due to the diffraction limit. Since F and D are related by the numerical aperture, $\Delta\theta_\delta/\Delta\theta_\lambda = N.A.\delta/\lambda$. Thus, for a given focal plane, the angular resolution is inversely proportional to F, meaning that thicker systems have better angular resolution.

If we could reduce $\delta$ by an order of magnitude, we could reduce F by an order of magnitude while still maintaining $\Delta\theta_\delta = \delta/F$. Unfortunately, it is not possible to reasonably reduce $\delta$ in the focal plane. It is possible, however, to effectively reduce $\delta$ through compressive coding. Coding consists of nonlocally remapping image fields with wavelength-scale 3D focal plane optics.

In a conventional imaging system, the focal plane averages wavelength-scale features within each pixel. The reference structure layer in the proposed system remaps the field in the focal plane such that wavelength-scale features from disjoint pixels are measured by each pixel. A pixel measurement in a conventional system may be modeled as

$$m = \int_A I(r)\,dr$$

where A is the area of the pixel. With compressive coding, the ith pixel measurement is

$$m_i = \int h_i(r)I(r)\,dr$$

where $h_i(r)$ is a non-local map of the focal intensity onto the ith pixel. $h_i(r)$ is a non-convex distribution. Focal plane coding consists of selecting $h_i(r)$. Focal plane coding may be used to improve resolution or to produce "thin" imaging systems for the visible and infrared spectral ranges.

## 2. COMPRESSIVE SAMPLING

We consider a multiple aperture imaging system where each aperture includes an imaging lens, a focal plane coding element and an electronic focal plane. Image synthesis from multiple aperture systems was pioneered in the TOMBO system [1, 2]. The lens-focal length distance is adjusted such that an array of images is formed on the focal plane. We assume for simplicity that the redundant images are identical, although aperture to aperture variations in sampling may be corrected algorithmically and may be deliberately designed into conformal aperture arrays.

The focal plane intensity distribution formed by an aperture is $I(\mathbf{r}) = \int I_o(\mathbf{r}')h(\mathbf{r}-\mathbf{r}')\,d\mathbf{r}'$, where $I_o(\mathbf{r}')$ is the "true" intensity distribution of the object in its native space. The focal plane intensity distribution is integrated on the $i^{th}$ pixel of the $j^{th}$ aperture to obtain the measurement value

$$\begin{aligned} m_{ij} &= \int p_{ij}(\mathbf{r})I(\mathbf{r})\,d\mathbf{r} \\ &= \iint I_o(\mathbf{r}')p_{ij}(\mathbf{r})h(\mathbf{r}-\mathbf{r}')\,d\mathbf{r}'d\mathbf{r} \end{aligned} \qquad (1)$$

where $p_{ij}(\mathbf{r})$ is the focal plane code for the $i^{th}$ pixel in the $j^{th}$ aperture. If we expand the object intensity using for example a sinc or wavelet basis, $I_o(\mathbf{r'}) = \sum_n s_n \psi_n(\mathbf{r'})$ then Eqn. (1) becomes

$$m_{ij} = \sum_n H_{ijn} s_n \tag{2}$$

where $m$ and $s$ are measurement and object state vectors, respectively, and

$$H_{ijn} == \int\int \psi_n(\mathbf{r'}) p_{ij}(\mathbf{r}) h(\mathbf{r}-\mathbf{r'}) d\mathbf{r'} d\mathbf{r} \tag{3}$$

As a first approximation, one may assume a local object basis such that the coefficients $s_n$ correspond to high resolution pixel states in the image. If we break the image into sub-blocks as in the original JPEG standard, for example, we may consider the linear transformation $\mathbf{m} = \mathbf{Hs}$ as implemented on each sub-block. For a $4 \times 4$ sub-block size, for example, the object state vector $\mathbf{s}$ corresponds to optical resolution cells as shown here:

$$
\begin{array}{cccc}
s_1 & s_2 & s_3 & s_4 \\
s_5 & s_6 & s_7 & s_8 \\
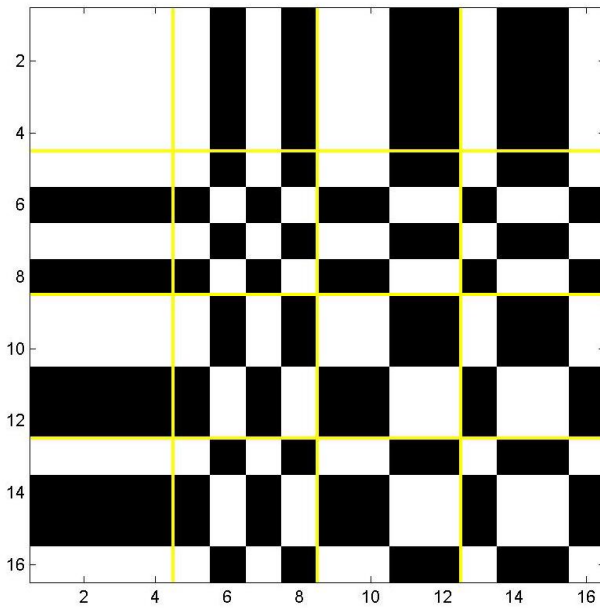s_9 & s_{10} & s_{11} & s_{12} \\
s_{13} & s_{14} & s_{15} & s_{16}
\end{array}
$$



Figure 1: Mask patterns for 4x4 Hadamard blocks (Spatial-code association is not unique).

| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | -1 | -1 | -1 |
| 1 | 0 | 0 | -1 | -1 | 0 | 0 | 1 |
| 1 | 0 | -1 | -1 | 1 | 1 | 0 | -1 |
| 1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 |
| 1 | -1 | 0 | 1 | -1 | 0 | 1 | -1 |
| 0 | -1 | 1 | 0 | 0 | 1 | -1 | 0 |
| 0 | -1 | 1 | -1 | 1 | -1 | 1 | 0 |

Figure 2: 8x8 transformation code for quantized cosine transform.

A mapping is implemented on $\mathbf{s}$ by masking the block in each subaperture with a different focal plane code. The entire block is integrated on a single pixel in each subaperture. In the case of $4 \times 4$ sub-blocks, $\mathbf{H}$ is a $16 \times 16$ matrix and $\mathbf{m}$ is a $16 \times 1$ vector. For example, if $\mathbf{H}$ is a Hadamard-S matrix (which is a 0-1 matrix obtained by shifting the Hadamard matrix elementwise up by 1 and scaling it by ½), the masks or codes for the 16 subapertures are as shown in Fig. 1. The figure shows the optical transmission pattern over each square pixel on the focal plane (white=1, black =0). Each pixel is segmented into a 4x4 grid of optical resolution elements. In the first subaperture, each electronic pixel integrates all incident optical power according to the code (all 1s) in the upper left corner block. In the second subaperture, the second and fourth columns of the source distribution $\mathbf{s}$ are blocked according to the code in the (1, 2)

block in the transmission pattern.  A complete image is acquired using 16 sub-apertures, each following a specific code as described.

We also consider the case where the elements of **H** are drawn from the set (-1, 0,1). In this case, $\mathbf{H} = \mathbf{H_1} - \mathbf{H_2}$, both $\mathbf{H_1}$ and $\mathbf{H_2}$ draw elements from the binary set (0,1). Coding schemes based on such matrices can be implemented easily. We illustrate next the compressive design of the transformation matrices and image reconstruction. The non-compressive design may be viewed as an extreme case where all measurements are used.



| PHT | DCT | QCT |



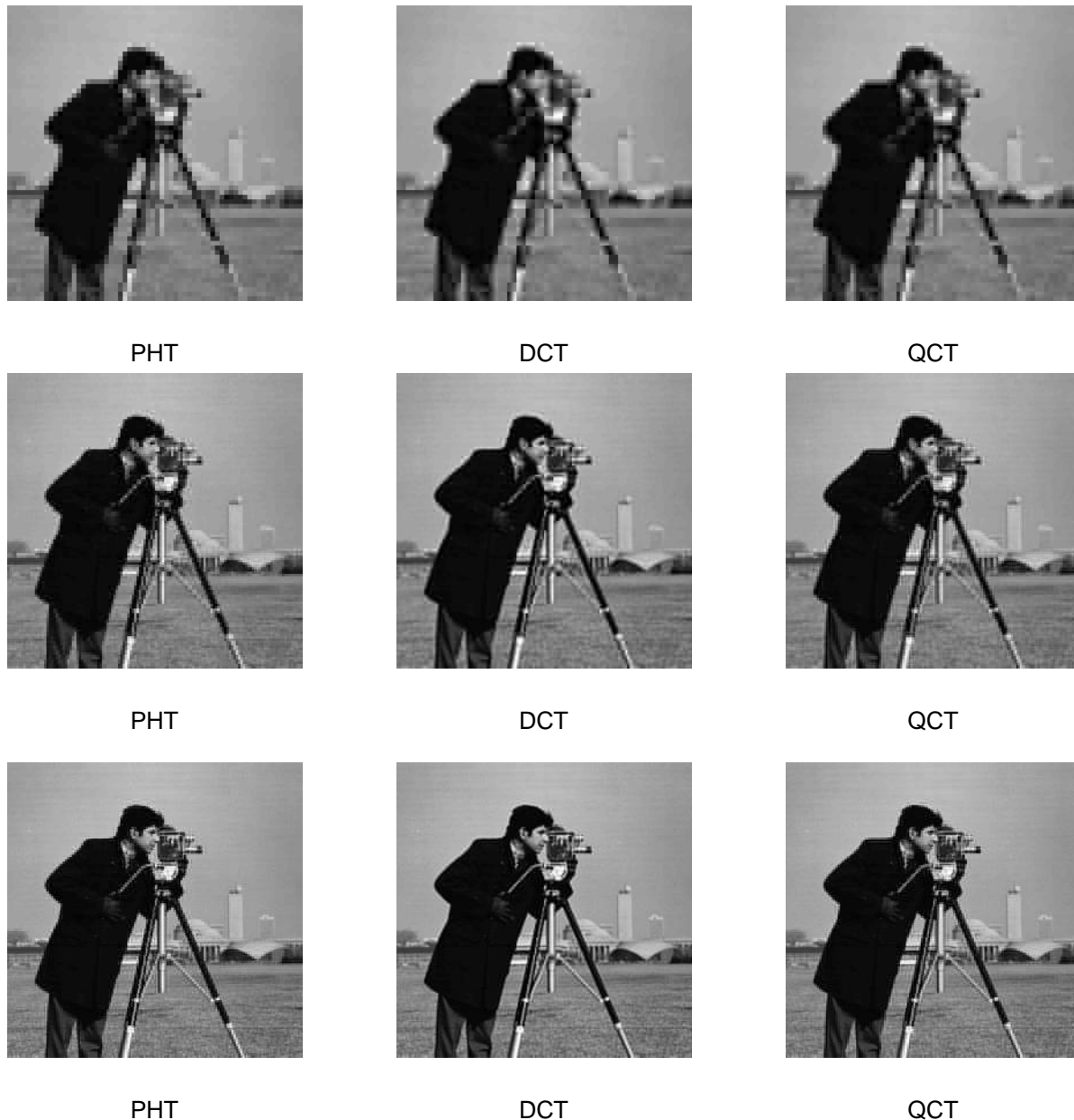| PHT | DCT | QCT |



| PHT | DCT | QCT |

Figure 3: Reconstructions using 4.69%,  15.63% and 32.81% of transformed components/available measurements.

In compressive system design, we use certain transforms to enable measurements of the principal components of the source image in a representation and source estimates by numerical decompression with high fidelity.  We introduce a couple of such transforms, which are new to our knowledge. Partition an image source into blocks of, for example, 8x8

pixels. Consider the two-dimensional transformation of each 8x8 block $\mathbf{S_{ij}}$ , $\mathbf{C_{ij}} = \mathbf{Q}\,\mathbf{S_{ij}}\,\mathbf{Q^T}$, where the transform matrix $\mathbf{Q}$ is defined as in Fig. 2.

The transform matrix has the following properties. Its elements are from the set (0, 1, -1), implying that the transform can be easily implemented as a mask. The rows of the matrix are orthogonal. The row vectors are quite even in Euclidean length, with the ratio $\sqrt{2}$ between the largest and the smallest. When the source image is spatially continuous within block $\mathbf{S_{ij}}$, the transformed block $\mathbf{C_{ij}}$ exhibits the compressible property that its elements decay along the diagonals. We may therefore truncate the elements on the lower anti-diagonals and measure only the remaining elements with fewer sensors. Denote by $\overline{C_{ij}}$ the truncated block matrix. We then get an estimate of the source block $\mathbf{S_{ij}}$ from $\mathbf{Q^{-1}}\ \overline{C_{ij}}\ \mathbf{Q^{-T}}$ (decompression). The same transform matrix is used for all blocks of image $\mathbf{S}$.

The above ideas are similar to the image compression with the discrete cosine transforms, as used in the JPEG protocol. In fact, the specific matrix $\mathbf{Q}$ can be obtained by rounding the discrete cosine transform (DCT) of the second kind into the set (0, 1, -1). We therefore refer to $\mathbf{Q}$ as the *quantized* cosine transform (QCT). But the very structure of the QCT matrix itself can be used to explain the compression.

Simulation results for QCT sampling and reconstruction are provided in Fig. 3. Visually the effectiveness of the compression with the QCT is surprisingly close to that with the DCT. We also use a permuted Hadamard transform (PHT) with row ordering [1, 5, 7, 3, 4, 8, 2, 6]. We skip here the quantitative comparisons among these transform. Based on the basic 8x8 QCT and PHT matrices, we can also construct larger transform matrices of hierarchical structure for multiple resolution analysis.

## 3. FOCAL PLANE CODING MASKS

We have previously considered transmission masks in coherent imaging and interferometry as field sampling elements displaced above the focal plane [3]. In the present context we consider sampling masks directly in contact with the focal plane. As a first step to demonstrating focal plane coding using transmission masks, we have experimentally shown that sub-pixel apertures can create sub-pixel response on the CCD sensors. We have also used these masks to characterize point-spread function of the imaging system by using the sub-pixel aperture scanning technique.

Most imaging systems are linear imaging systems. In a linear imaging system, the image $i(x, y)$ and the object $s(x, y, \lambda)$ are related as

$$i(x,y) = \int_{\lambda} d\lambda \iint_{x',y'} s(x',y',\lambda)h(x,y;x',y';\lambda)dx'dy' , \qquad (4)$$

where $h(x,y;x',y';\lambda)$ is the impulse response function, also called point spread function (PSF) at the wavelength $\lambda$. In a digital electronic imaging system, the image is sampled by the photo-detector array. The signal from the pixel of (m, n) is

$$i(m,n) = \iint i(x,y)p(x-ma,y-na)dxdy , \qquad (5)$$

where $p(x,y)$ is the pixel response function

$$p(x,y) = \begin{cases} 1 & \text{for } |x| \le a/2 \text{ and } |y| \le a/2 \\ 0 & \text{otherwise} \end{cases} , \qquad (6)$$

Rewriting Eqn. (5), the signal from the (m, n) pixel is

$$i(m,n) = \iint_{x,y} p(x-ma,y-na)dxdy \int_{\lambda} d\lambda \iint_{x',y'} s(x',y',\lambda)f(x,y;x',y';\lambda)dx'dy' \qquad (7)$$

The pixel size of the CCD is 5.6 micron square with an individual micro-lens on the top of each pixel. The CCD has a total number of 650 x 490 pixels. The sub-pixel mask is a 120 nm chrome mask on a glass substrate. One mask pattern

is shown in Fig. 4. There are four sub-pixel line apertures with one, two, three, and four micron width in the mask pattern. The pitch (center to center distance) of the sub-pixel apertures on the mask matches the pitch of the CCD pixels. We align the sub-pixel apertures to the pixels of the CCD. To show that this sub-pixel mask can create localized sub-pixel response, we used a high-NA (NA=0.5) objective lens to focus a far-field point source light on the mask. The far-field point source was created by focusing a HeNe laser (0.6328 nm wavelength) to a 15 micron pin hole. The polarization of the light is parallel to the line apertures (vertically in Fig. 4). Theoretically, the Airy radius of the focused spot size is 0.78 micron.
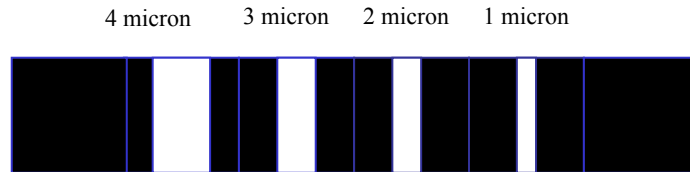


Fig. 4 The sub-pixel mask pattern

In the experiment, the mask coded CCD was on a high resolution translation stage. The movement resolution is about 50 nm. We translated the CCD to scan the focused spot crossing the sub-pixel apertures with an increment step of 0.2 micron. Figure 5 shows the signals from four mask coded pixels versus the scanning distance. The vertical axis is the pixel number by the size of the line aperture. Figure 6 plots the signals of four pixels for scanning distance. The signals from pixels with three and four micron apertures have flattops. This indicates the spot size is smaller than the sizes of the apertures. Therefore, the pixels with three and four micron apertures capture all the light. When the focused spot (PSF) overlapped with the two micron and one micron apertures, only partial light was transmitted.
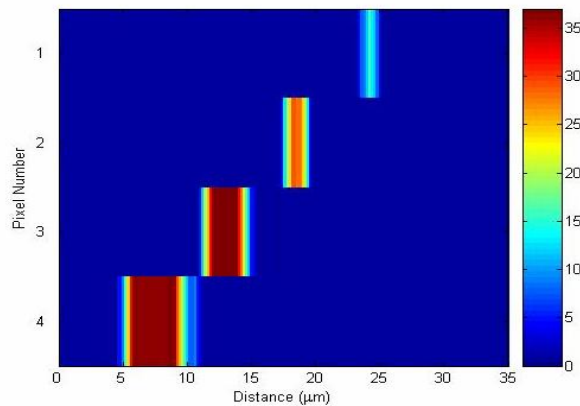


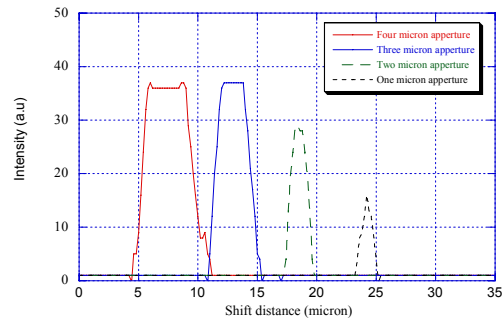Fig. 5. Signals from coded aperture pixels versus the scanning distance.

Fig. 6.  Signals from coded aperture pixels versus the scanning distance.

In summary, we have shown that the sub-pixel mask can create localized response in each individual pixel of the CCD camera.

## ACKNOWLEDGEMENT

## REFERENCES

1.  Tanida, J., et al., *Color imaging with an integrated compound imaging system.* Optics Express, 2003. 11(18): p. 2109-2117.
2.  Tanida, J., et al., *Thin observation module by bound optics (TOMBO): concept and experimental verification.* Applied Optics, 2001. 40(11): p. 1806-1813.
3.  Tumbar, R. and D.J. Brady, *Sampling field sensor with anisotropic fan-out.* Applied Optics, 2002. 41(31): p. 6621-6636.