Super-Resolution Image Reconstruction for Diverse Sub-Imager Arrays
— Source link ↗

Sally L. Wood, Hsueh-ban Lan, Dinesh Rajan, Marc P. Christensen

Institutions: Santa Clara University, Southern Methodist University

Topics: Image resolution, Iterative reconstruction and Image sensor

Related papers:

- Improved Multiplexed Image Reconstruction Performance Through Optical System Diversity Design

- A Hardware Realization of Superresolution Combining Random Coding and Blurring

- Three-dimensional super-resolution: theory, modeling, and field test results

- Electro-optical foveated imaging and tracking system

- Compressive coded apertures for high-resolution imaging

Super-Resolution Image Reconstruction for Diverse Sub-Imager Arrays

Sally L. Wood   Hsueh-Ban Lan
Santa Clara University
500 El Camino Real
Santa Clara, CA 95053 USA

Dinesh Rajan   Marc P. Christensen
Southern Methodist University
6251 Airline Road
Dallas, TX, 75275

Abstract- **High resolution images can be estimated using multiple low resolution images obtained from an array of sub-imagers with overlapping fields of view. Design choices for the optics and sensors of a flat camera can have a significant impact on the performance of reconstruction algorithms. This paper will analyze designs that provide diversity which reduces the expected error of reconstruction algorithms. It will be assumed that the fields of view of individual sub-imagers can be controlled so that the desired resolution improvement at the object distance of interest can be achieved. Simulations will demonstrate the analytical results.**

I. INTRODUCTION

A thin modular flat form factor camera could create numerous new imaging application areas for which a traditional camera is too bulky, fragile, or conspicuous. The boxy form factor of a camera with a lens system and a high resolution image sensor is determined by the parameters of the optical elements. One way to reduce the camera thickness is to use optical elements that can be positioned closer to the image sensor while maintaining the same point spread function. This will project a smaller image onto the sensor, and since the size of the individual sensor elements can not be reduced proportionately, the captured digital image will have a lower resolution limited by the detector size, not the optics. However, if the imaging system has a compound eye design which produces multiple low resolution images with overlapping fields of view, computational imaging techniques can be used to reconstruct a higher resolution image.

Several new computational imaging applications have been reported in recent years [1-9] using methods similar to those used in radio astronomy [10] and medical image reconstruction applications [11]. If a moving low resolution camera provides a sequence of images of a stationary scene from multiple shifted perspectives, image registration and reconstruction algorithms can be used to create a higher resolution image that looks the same as an image captured by a more expensive camera [2]. Parallel acquisition of multiple images can use an array of sub-imagers such as the TOMBO designs [3-5]. Each sub-imager consists of a micro-lens and a small pixel sensor array. With fixed arrays and optics, the registration problem is greatly reduced but only one specific object distance will produce the correct overlap of fields of view.

The Panoptes approach to flat camera design is different from these two applications in several respects [6-9]. The high resolution images produced by the system will be used by automated image analysis systems, so the measure of performance is the information captured in the reconstruction rather than measures based on human perception. Like the TOMBO system, the low resolution images are acquired in parallel from an array of sub-imagers, but active control of the look direction of individual sub-imagers through micro-mirror arrays allows variable resolution improvement and variable object range. In addition, diversity in the sub-imager array design is used to reduce the reconstruction error by eliminating the null space due to a uniform sensor array [6, 7, 9].

II. IMAGE ACQUISITION MODEL

Let f(x,y) be a continuous two-dimensional light intensity projected onto the plane of detector elements. Let p(x,y) be the continuous response of an individual pixel detector in the detector array. Then $g(x_0,y_0)$, the response of an individual pixel detector centered at (x_0,y_0), is given in (1).

$$g(x_0, y_0) = \iint p(x - x_0, y - y_0) f(x, y)\, dx\, dy \qquad (1)$$

For an array of identical detectors with detector spacings of Δx and Δy, the complete acquired image $g(m_x,m_y)$ can be represented as the convolution g(x,y) = f(x,y)*p(-x,-y) sampled at intervals of Δx in x and Δy in y. The image data can be stored by rows in an M×1 vector, g.

The desired high resolution image, $f(n_x,n_y)$, which is f(x,y) sampled at intervals of $\Delta x/q$ and $\Delta y/q$, can be stored by rows in an N×1 vector, f, where N=q^2M. A linear relationship between f and g is given in (2) where the M×N observation partial matrix, **H**, may be space variant, and v is an M×1 vector of measurement noise values.

$$g = Hf + v \qquad (2)$$

Algorithms to restore or reconstruct f from g must address three problems. For most images undersampling by a factor of q will cause aliasing. In addition, the integrated response of an individual detector, p(x,y), will blur the desired image and heavily attenuate some spatial frequencies. This will result in an ill-conditioned estimation problem. It is assumed that the point spread function of the optics is much smaller

than the individual pixel detector extent, so optical blur will make little or no contribution to blur at the desired resolution. And finally, the measurement noise will cause errors in the estimate of f, and the impact of the noise on the errors is increased when H is poorly conditioned.

If multiple shifted low resolution images can be acquired, the aliasing of g relative to f can be eliminated. Assume that the objects in the desired image are far enough away so that a small lateral shift of perspective simply shifts the whole projected image, and at the desired resolution there is no measurable disparity due to object range differences. Let the $(k,l)^{th}$ image of a set of K×L low resolution images be given by $g_{k,l} = H_{k,l}f + v_{k,l}$. All the sub-imager observations can be combined into a single vector as shown in (3), and the combined **H** matrix and observations, g, will have the same form as (2). If the shifts are $\Delta x/q$ and $\Delta y/q$, and a $q \times q$ array of shifted low resolution images is acquired from q^2 sub-imagers, then effectively g(x,y) has been sampled at the desired sampling interval for f(x,y) [12], and a simple reordering of **H** and g would represent that.

$$\mathbf{g} = [\ \mathbf{g}^T_{00}\ \ \mathbf{g}^T_{01} \cdots\ \mathbf{g}^T_{0\,L-1}\ \ \mathbf{g}^T_{10}\ \cdots\ \mathbf{g}^T_{K-1\,L-1}\ ]^T \quad (3)$$

$$= [\mathbf{H}^T_{00}\ \ \mathbf{H}^T_{01}\ ...\ \mathbf{H}^T_{K-1\,L-1}\ ]^T\ \mathbf{f}\ + \mathbf{v} = \mathbf{Hf} + \mathbf{v}$$

After an array of sub-imagers has acquired enough image data to eliminate the aliasing problem, a minimum variance of error estimator (MVE) can be used to estimate f from the blurred and noisy observations in g using (2) or (3). The reconstruction, $\hat{\mathbf{f}}$, is given in (4) where $f_0 = E\{f\}$ for the class of images to be reconstructed. P_0 is the covariance of the image class, and R_v is the covariance of the noise. An alternative form is shown in (5). In both (4) and (5) the actual value for the covariances is not known exactly, so estimates are used. The noise, v, is assumed to be zero-mean and uncorrelated with the image.

$$\hat{f} = f_0 + \hat{P_0}\,H^T\left(H\,\hat{P_0}\,H^T + \hat{R_v}\right)^{-1}\left(g - Hf_0\right) \quad (4)$$

$$\hat{f} = f_0 + A\left(g - Hf_0\right)$$

$$\hat{f} = \left(H^T\,\hat{R_v}^{-1}\,H + \hat{P_0}^{-1}\right)^{-1}\left(H^T\,\hat{R_v}^{-1}\,g + \hat{P_0}^{-1}\,f_0\right) \quad (5)$$

For this estimate, the expected mean squared error (emse) can be computed from the average of the trace of ξ in (7).

$$\tilde{f} = f - \hat{f} \quad (6)$$

$$\xi = E\left\{\tilde{f}\left(\tilde{f}\right)^T\right\} = (I - AH)P_0(I - AH)^T + AR_vA^T \quad (7)$$

The singular value decomposition of **H** can be used to determine how poorly conditioned the observation set is. In the singular value decomposition $\mathbf{H} = \mathbf{USV^T}$, the M×N diagonal matrix **S** holds the singular values, and **U** and **V** are M×M and N×N square unitary matrices. If $P_0 = p_0I_{N\times N}$ and $R_v = \sigma^2 pI_{M\times M}$, $\mathbf{\xi} = \mathbf{V\Lambda V^T}$ where the elements of the diagonal matrix, $\mathbf{\Lambda}$, are given by (8). When $s_{ii} = 0$, $\lambda_{ii} = p_0$.

$$\lambda_{ii} = \frac{p_0\sigma^2}{p_0 s_{ii}^2 + \sigma^2} \quad (8)$$

A frequency domain perspective can explain the presence of very small singular values even when sufficient sub-imagers are used to avoid aliasing. If the continuous p(x,y) is modeled as a uniform average over a square detector surface of width a, the spatial frequency response, P(u,v), is given in (9). Any spatial frequency that is an integer multiple of (1/a) will have a frequency response of zero. For spatial frequencies which have a very low response due to the detector size and shape, noise amplification is a significant problem when high gains are used to restore these frequencies in the reconstructed image.

$$p(x,y) = \frac{1}{a^2}\text{rect}\left(\frac{x}{a}\right)\text{rect}\left(\frac{y}{a}\right)$$
$$P(u,v) = \text{sinc}\,(au\,)\text{sinc}\,(av\,) \quad (9)$$

III. ADDING DIVERSITY

Adding additional sub-imagers with different detector sizes, different magnifications, or different rotations can provide information about spatial frequencies that are lost with a uniform array of sub-imagers. These diverse sub-imager observations can be combined in (10) in the same way that was used in (3) for different uniform sub-imager arrays. Then previously discussed reconstruction methods and error analysis can be used. For this example with three types of sub-imager arrays, (11) can be used to compute the singular values.

$$g = \begin{bmatrix} g_I \\ g_{II} \\ g_{III} \end{bmatrix} = \begin{bmatrix} H_I \\ H_{II} \\ H_{III} \end{bmatrix} f + v \quad (10)$$

$$H^T R_v^{-1} H = \quad (11)$$
$$H_I{}^T R_{v_I}^{-1} H_I + H_{II}{}^T R_{v_{II}}^{-1} H_{II} + H_{III}{}^T R_{v_{III}}^{-1} H_{III}$$

Fig. 1 compares the spatial distribution of expected errors for reconstruction of a small 35×35 image tile of f with q = 3. A uniform magnification system on the left has 27 sub-imagers with $a=3d$, where d is the desired resolution. A diverse system on the right has $a=3d$, 4d, and 5d and uses the same number of observed pixels. All four plots are scaled separately with white representing the lowest expected error

and black representing the highest. The upper row shows the emse for the full tile. For the diverse system the highest errors are on the edges while for the uniform system the highest errors are much larger and have a periodic pattern. Trimming the tiles to 31×31, shown in the lower row, reduces the maximum error for the diverse system, but not the uniform system. Thus, for the diverse system, individual small tiles of a large image can be successfully reconstructed independently. If trimmed tiles are used, the tile edges must overlap.
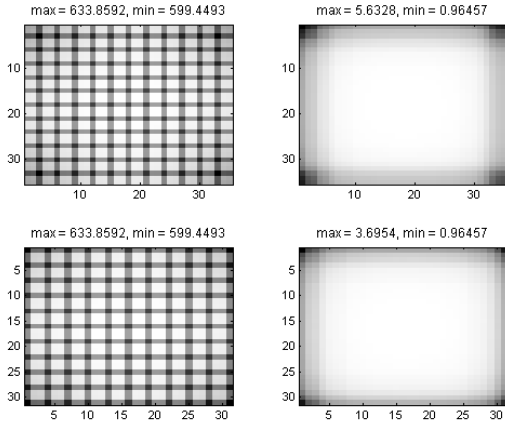

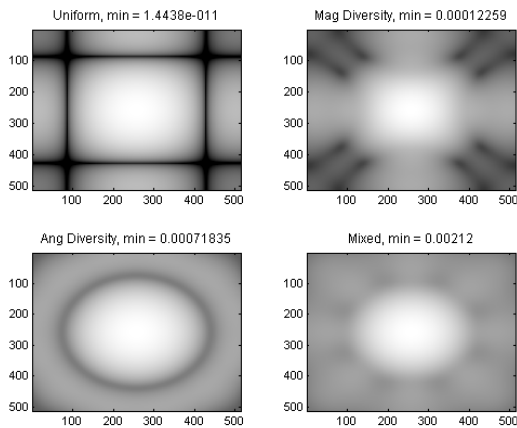Fig. 1. Spatial distribution of expected squared errors for full 35×35 tile and trimmed 31×31 tile.

Fig. 2. Combined frequency responses for four systems.

The distribution of errors also can be evaluated as a function of spatial frequency. Fig. 2 shows the sum of the squared frequency response of the continuous p(x,y) responses as a continuous function of (u,v) for four systems with q= 3. All four systems use the same scaling with black representing low values of the sum, which would cause a high expected squared error in the estimate of F(u,v). The upper left shows the response for a uniform system, and the upper right shows the response of a system with magnification diversity using a = 3d, 4d, and 5d. In the lower

left there is uniform magnification a = 3d with three sets of sub-imager arrays rotated by 0, π/6 and π/3. The result in the lower right uses four sub-imager types with a = 3d and 4d and rotations of 0 and π/4. This has the highest value for the minimum average sum which will correspond to the lowest value for the expected error in the estimate of F(u,v).

IV. SIMULATION RESULTS

The emse values for a 35×35 pixel image reconstruction with q=3 are plotted in Fig. 3 as a function of actual measurement noise variance, σ^2, for five array architectures. In all five cases the same number of observations was used. For each architecture four estimators using different values for the expected noise variance were tested. The plots shown with dashed lines represent uniform arrays with a=3d. High attenuation of some spatial frequencies causes all four estimators for this architecture to produce virtually the same results. Reducing the actual noise variance below 1, does not result in any expected performance gain.
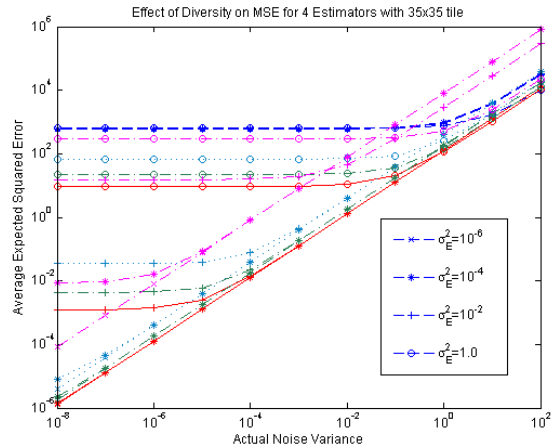

Fig. 3. Expected squared error for a 35x35 reconstruction tile with q = 3 comparing magnification diversity.

It has been assumed that the observations and estimated pixel values are scaled to the range of [0,256) to be consistent with typically used 8-bit per pixel image data. However, it is not assumed that observations are limited to 8 bits. An increased number of bits will be needed to represent the lower measurement noise variance observations, and more than 8 bits per pixel will be needed to represent the information in the images used for automatic processing rather than for human observation.

Solid line plots show a diverse system where some of the sensors with a=3d have been replaced with sensors with a = 4d or 5d. The larger sensors have different spatial frequency responses and the combined system provides improved performance. Reducing the actual noise variance reduces the emse of the reconstruction until it is limited by the expected noise variance used by the specific estimator. Similar results are obtained for a = 3.5d and 4d, shown with a dash-dot

lines. The worst performer of the four diverse systems has a mild diversity with $a = 3.1d$ and $3.2d$. It performs better than the homogeneous system and it still allows estimators to benefit from a reduction in the measurement noise.

Fig. 4 shows the singular values of the combined **H** matrix for each of the five architectures used in Fig. 3. The uniform system shown with the dashed line has higher values over most of the range, but then has a large number of values that are zero. The diverse system shown with the solid line has lower values at the left, but has the highest minimum value.
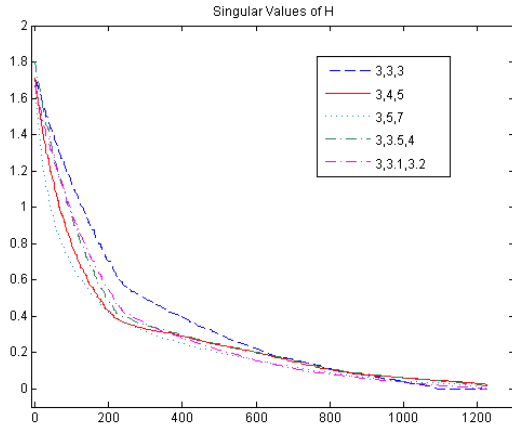

Fig. 4. Singular values for **H** matrices for systems in Fig. 3.

A. Effect of Tile Size

Fig. 5 compares the performance of uniform and diverse systems when the linear dimension of the tile size is doubled. Although the larger tile shows a small improvement over the smaller one, the small tile of the diverse system is far superior to the large tile of the uniform system. This indicates that diverse systems can reduce computational complexity by independently reconstructing very small sections of a larger image. This allows parallel computation of multiple independent small images tile reconstructions which is also computationally advantageous.

B. Effect of Resolution Improvement

Fig. 6 compares the emse for a 41×41 pixel image tile with q = 3, 5, and 7. As the desired resolution improvement increases, the noise level of the measurements must be reduced to maintain the same noise level in the reconstructed image. For the two uniform systems with q = 3 and 5, the emse does not decrease as the measurement noise variance is reduced below 0.1, and the expected errors are higher for the resolution improvement of 5. For the diverse systems, the emse plots are almost exactly parallel and the increase in error level as q increases from 3 to 5 to 7 is easily observed.
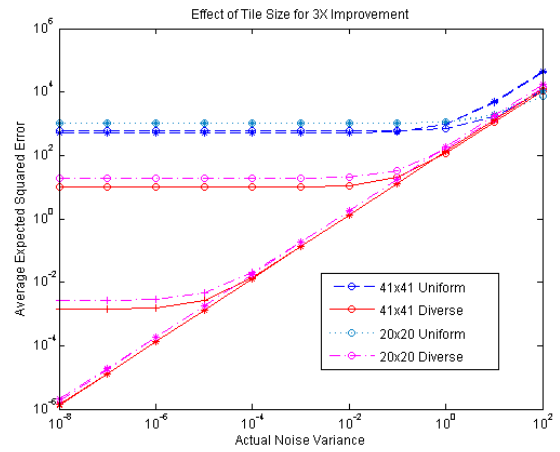

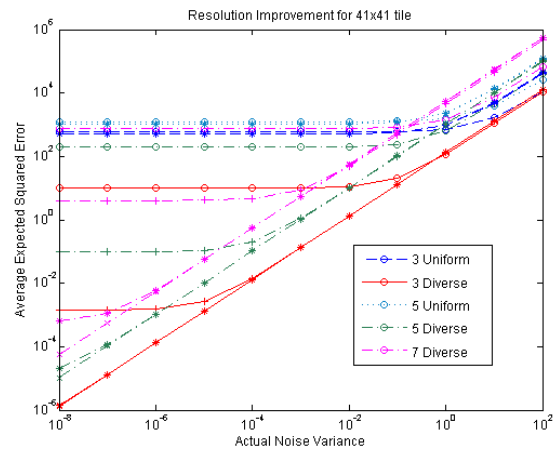Fig. 5. Tiles sizes of 400 pixels and 1681 pixels.

Fig. 6. Resolution improvement of q = 3, 5, and 7 for a 41×41 pixel image tile.

Diverse systems have been compared using the emse as a performance measure, but other measures of quality may also be appropriate. For images that will provide input for automatic analysis and detection systems, the overall performance of the complete system should be used to compare sub-imager architectures. In [9] the mean squared error for reconstructed pixel values was shown to have a similar behavior to the mean squared error of the edge gradient for several actual images, synthetic images, and noise images.

In Fig. 7 a low contrast image of the Bay Bridge taken from the USC image data base is reconstructed with q = 3 and σ^2=0.01 using 27 sub-imagers. Results for a uniform and diverse system are compared. For the diverse reconstruction in Fig. 7d, the edge detail is a good match to the original image tile in Fig. 7b. Noise amplification and periodic structure are apparent in the uniform reconstruction in Fig. 7c.
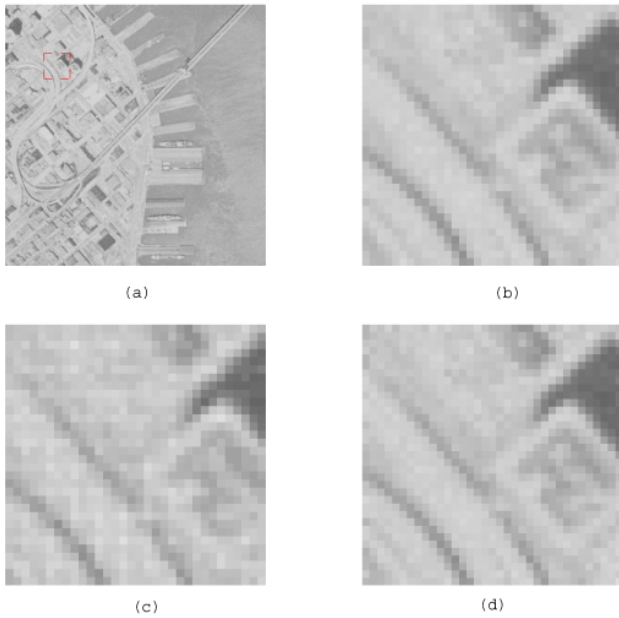
Fig. 7. USC image data base Bay Bridge image (a) and a 35×35 image tile (b). Reconstruction with q=3 for uniform (c) and diverse (d) systems.

IV. CONCLUSIONS

For a flat form factor camera design based on arrays of low resolution sub-imagers, an architecture using magnification and orientation diversity has several demonstrated advantages. The diversity allows information about all spatial frequencies to be captured by some group of sub-imagers. Diverse systems can be compared in terms of spatial error distribution or error distribution in spatial frequency.

Simulations of several diverse architectures show that even a small amount of diversity improves the performance as measured by expected mean squared error [6,7]. Similar performance was achieved for a variety of systems, and the improvement does not rely on any precise relationship between the sub-imagers parameters. This allows flexibility for calibration and adaptive control of resolution improvement using steerable micro-mirror arrays

Since small tiles of a larger image can be reconstructed independently using a diverse system, local computation can be done efficiently using exact estimators without assumptions of space invariance or circulant approximations. Reduction in the expected error from increased tile size was much smaller than the error reduction due to added diversity. This structure can also use parallel computation for the independent reconstruction of small image tiles.

Although resolution improvement can be controlled by the amount of overlap of individual detector fields of view, the impact of measurement noise on the reconstructed image increases with the desired resolution improvement. This implies that when the desired resolution improvement is increased, the measurement noise must be reduced to achieve

the same error rate and the number of significant bits in the measurements must increase.

REFERENCES

[1] S. Chaudhuri, ed. *Super-Resolution Imaging*, Kluwer Academic Publishers, Boston MA, 2001.

[2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Multiframe Super-resolution," IEEE Transactions on Image Processing, vol 13, no. 10, pp1327-1344, October, 2004.

[3] J. Tanida, et al., "Thin Observation Module by Bound Optics (TOMBO): Concept and experimental verification," *Appl. Optics-IP*, vol. 40, Issue 11, p. 1806-1813, Apr. 2001.

[4] Y. Kitamura "Reconstruction of a High-Resolution Image on a Compound-Eye Image-Capturing System," *Appl. Optics*, vol. 43, Issue 8, p. 1719-1727, Mar. 2004.

[5] K. Nitta, R. Shogenji, S. Miyatake, J. Tanida, "Image reconstruction for thin observation module by bound optics by using the iterative backprojection method," *Appl. Optics*, Vol. 45, Issue 13, p. 2893-2900, May. 2006.

[6] S. L. Wood, B. J. Smithson, D. Rajan, and M. P. Christensen, "Performance of a MVE Algorithm for Compound Eye Image Reconstruction Using Lens Diversity," Proc IEEE ICASSP 2005, Vol. II, pp.593-596, Philadelphia Pa, March 18-23, 2005.

[7] H-B. Lan, S. L. Wood, M. P. Christensen, D. Rajan, "Benefits of optical system diversity for multiplexed image reconstruction," *Appl. Optics*, Vol. 45, Issue 13, pp. 2859-2870, May, 2006.

[8] M. P. Christensen, V. Bhakta, D. Rajan, S. C. Douglas, S. L. Wood, M. W. Haney, ``Adaptive flat multi-resolution multiplexed computational imaging architecture utilizing micro-mirror arrays to steer sub-imager field-of-views,'' *Appl. Optics*, Vol. 45, Issue 13, pp. 2884-2892. May 2006.

[9] S. L. Wood, H-B Lan, D. Rajan, M. P. Christensen, "Improved Multiplexed Image Reconstruction Performance Through Optical System Diversity Design," Proc IEEE ICIP, Atlanta, Georgia, Oct 8-11, 2006.

[10] R. N. Bracewell, *Two-Dimensional Imaging*, Prentice Hall, Englewood Cliffs, N.J., 1995.

[11] A. Macovski, *Medical Imaging Systems*, Prentice Hall, Englewood Cliffs, N.J., 1983.

[12] S. Mitra, *Digital Signal Processing* 3rd ed., Prentice Hall, 2006.