Reconstruction Algorithms for Compound Eye Images Using Lens Diversity

Sally L. Wood
Department of Electrical Engineering
Santa Clara University
Santa Clara, California, 95053 USA

Dinesh Rajan *Marc P. Christensen*
Department of Electrical Engineering
Southern Methodist University
Dallas Texas, 72275 USA

***Abstract* - A traditional optical imaging device uses a single high quality lens designed to project the field of view onto a sensor array for direct image creation. A camera with a flatter form factor using an array of smaller lens elements would create an array of smaller lower resolution images. Reconstruction of high resolution images from multiple low resolution images at various displacements is a well studied, ill posed problem. Using lenses with different imaging characteristics improves theoretical performance results and also reduces the general image reconstruction problem to a set of loosely coupled smaller reconstructions. Performance limits for reconstruction from multiple lower resolution images is considered as a function of measurement bit precision, measurement noise, and lens model uncertainty.**

I. INTRODUCTION

Typical optical imaging devices have a single lens selected to focus the desired image onto a sensor array at a desired magnification. For a good quality system that is focused correctly, the digitized output of the sensor can be viewed directly without further processing. The major cost associated with this type of imaging system is usually the optical components. The length requirement of the optical column which will focus the image determines the form factor for the device, and for a traditional system, this is not compatible with a flat form factor. A flat camera, if it could be designed, would have many potential applications ranging from conformal sensing skins on surveillance aircraft to head-mounted camera patches for firemen and rescue workers.

In many applications, such as radio astronomy or medical imaging [1,8] for example, the traditional imaging approach is not used. Instead, integrated measurements of the image field are sampled with lower resolution elements, and image reconstruction methods are used to compute higher quality images. Recent advances in optical imaging sensors have created an opportunity to use similar image reconstruction methods in the design of imaging devices with a flat form factor [5]. For example, the TOMBO imaging device [7,9] uses an array of small lenses and CCD sensors to create a corresponding array of overlapping low-resolution images which are recombined to form a higher-resolution image. Other designs have been proposed [4,6,10,11] which would improve the quality of the reconstructed image, allow range and resolution control, and be computationally reasonable. This type of reconstruction task uses methods similar to those of computed tomography in medical imaging [5] and image registration [e.g. 2].

The problem of reconstructing high resolution images from multiple overlapping low resolution images is a well-known theoretical super-resolution problem [see e.g. 3]. This problem has many challenges in both theoretical analysis and in practical application with noisy measurements and quantization limits. The resolution of the reconstruction will depend on how many different low resolution images are available and the amount of relative translation between them at the range of the scene of interest. In theory, many small translations are needed for high resolution. But small translations also result in a small amount of new information in the data from each additional low resolution image. For meaningful computational results, smaller translations will require a higher bit precision in the measurements to represent small but significant differences and a larger number of bits to prevent overflow problems.

In this paper a measurement model is presented for the acquisition of an array of low resolution images using a micro-lens array. A reconstruction method is developed for this model, and the expected error is evaluated as a function of image size, resolution improvement, and noise level. The objective is to develop computationally efficient procedures for high-quality image reconstruction. Partitioning the problem into independent reconstructions of small image tiles can reduce computation, but all the

information in the measurements may not be used. A second level of computation would be needed to form the best complete image from the tiles.

This paper also explores the use of lens diversity to decouple the individual tile reconstructions and improve the quality of the final image. Lens diversity can be achieved by adding observations taken from sensor/lens systems with different magnifications or by using lenses with a space variant magnification. The solution with multiple lens magnifications is shown to lead to a modular reconstruction method that allows small loosely coupled sub-tiles to be estimated separately. A similar method can be used for the space variant magnification model. This technique allows parallel computation of the individual tile results and computationally reasonable merging of multiple tiles. Both expected error computations and simulated reconstruction examples demonstrate the advantages of this approach.

II. Measurement Model

Consider a two dimensional source image, **f** , and an observed image, **g**, each stored in a column vector in row order. For the traditional single lens, high resolution imager, **g** is related to **f** by an observation matrix **H** as shown in Equation (1).

$$\mathbf{g} = \mathbf{H}\,\mathbf{f} + \mathbf{v}. \tag{1}$$

The vector **v** represents measurement noise. For the high resolution imager, the observations in **g** are viewed as our estimate of the desired image **f.** The **H** matrix represents the model of the imaging system, and if there is no change in magnification, the ideal **H** would be an identity matrix. Typically the system is in focus, and the point spread function of the imaging system is assumed to be small compared to pixel size.

In applications, **H** may represent both magnification changes and a finite point spread function. For large point spread functions, an image restoration step may be used when statistics of the noise and image source class are also known. The resolution of **g** is limited by the size of the sensor pixels and lens magnification. It is also assumed that the exposure time is chosen to provide the desired dynamic range of measured pixel values from the sensors.

A. Array of Sub-Imagers

The lens just described requires a "cubic" or large form factor. For a flat camera form factor would need a smaller lens positioned closer to a smaller sensor array. The change in magnification factor would reduce the resolution of **g**. Equation (1) would still represent the general measurement the model, but the length of **g** would now be much smaller compared to the traditional system, and **H** would be modified according to the characteristics of the small lens.

From a single small image, very little improvement in resolution could be expected from traditional image restoration methods. But an array of small imaging systems, called sub-imagers, could provide information to compute a higher resolution estimate of **f** using image reconstruction techniques. Each sub-imager would have a micro-optical system and a small pixel sensor array.

Assume that the desired image at the desired resolution has K_f rows and L_f columns. The column vector **f** will have $N= K_f * L_f$ elements. Let $\mathbf{g_{00}}$ be a J element pixel array associated with a single sub-imager. If the resolution of $\mathbf{g_{00}}$ is much lower than **f**, then J will be much smaller than N. Assume that the position and orientation of the subimagers is controllable. A second identical sub-imager, $\mathbf{g_{01}}$, can be positioned so that its field of view is shifted slightly in the horizontal direction. This is represented by a shift of X pixels in the desired image vector, using a shift matrix **Z** with elements $z(k, l) = \delta(k + 1 - l)$. This is shown in Equation (2). Assume that **f** spans the field of view for all sub-imagers so that when a shift operation post multiplies $\mathbf{H_S}$, all elements lost in shifting are zero. Similarly, let $\mathbf{g_{10}}$ be another identical sub-imager with a vertically shifted field of view. This is also shown in Equation (2) as a larger shift, since the image array in **f** is stored by rows.

$$\begin{aligned}
\mathbf{g_{00}} &= \mathbf{H_S}\,\mathbf{f} + \mathbf{v_{00}} = \mathbf{H_{00}}\,\mathbf{f} + \mathbf{v_{00}}.\\
\mathbf{g_{01}} &= \mathbf{H_S}\mathbf{Z}^{X}\,\mathbf{f} + \mathbf{v_{01}} = \mathbf{H_{01}}\,\mathbf{f} + \mathbf{v_{10}}.\\
\mathbf{g_{10}} &= \mathbf{H_S}\mathbf{Z}^{Y}\,\mathbf{f} + \mathbf{v_{10}} = \mathbf{H_{10}}\,\mathbf{f} + \mathbf{v_{10}}.
\end{aligned} \tag{2}$$

When these small images are formed with correctly oriented micro-optical systems and then are used to create an estimate of the desired image, it can be considered a problem of image reconstruction. New observations from each new small image add new information about **f**. A K x L array of M=K*L sub-imagers will produce M low resolution images with a total of M*J image pixels. Using all the pixel arrays from all the sub imagers, a better estimate of the original image can be made, for example as described in [7,9]. The combined observations from the array of sub-imager sensors is shown in Equation (3). Here the concatenated vector of small low-resolution images does not represent the final image we wish to view.

$$\begin{aligned}
\mathbf{g} &= [\,\mathbf{g}^{t}_{00}\;\; \mathbf{g}^{t}_{01} \cdots\; \mathbf{g}^{t}_{0\,L-1}\;\; \mathbf{g}^{t}_{10}\;\; \cdots\;\; \mathbf{g}^{t}_{K-1\,L-1}\,]^{t}\\
&= [\mathbf{H}^{t}_{00}\;\; \mathbf{H}^{t}_{01}\;\; \ldots\;\; \mathbf{H}^{t}_{K-1\,L-1}\,]^{t}\,\mathbf{f} + \mathbf{v} = \mathbf{H}\mathbf{f} + \mathbf{v}
\end{aligned} \tag{3}$$

In Figure 1 a slice through an array of sub-imagers is shown schematically. The desired image at the desired resolution is shown at the top of the figure, and the pixel arrays of individual sub-imagers are shown at the bottom of the figure. The magnification factor of the small lenses causes many pixels at the top to be projected onto a single pixel at the bottom. In this figure perfect focus is assumed, and the overlapping fields of view of different sub-imagers are shown.

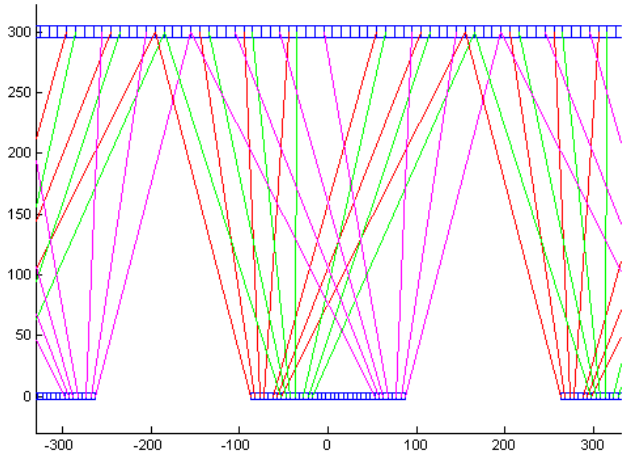

Figure 1: Slice through overlapping sub-array fields of view.

B. Sources of Noise

The noise vectors used in Equations (2-4) can be used to model uncorrelated measurement noise with a Gaussian distribution about a zero mean. It can also be used to model quantization noise as a uniformly distributed random variable between –w and w, where the value of w is determined by the number of bits used for each pixel value in **g**. Additional sources of noise with more complex statistics include random noise with a variance that depends on the values of **g** and noise due to model errors in **H**.

III. Image Reconstruction from Arrays

A linear minimum variance of error estimator, as shown in Equation (4), can be used to compute an estimate for **f** at the desired resolution using all of the observations in Equation (3). Here $\mathbf{f}_0$ is the expected average value for the image, $\mathbf{R}_\mathbf{v}$ is the expected covariance matrix for the noise vector **v**, and $\mathbf{P}_\mathbf{0}$ is the expected covariance matrix for the initial estimate of the image data.

$$\begin{aligned}\hat{\mathbf{f}} &= \mathbf{f}_0 + \mathbf{P}_\mathbf{0}\,\mathbf{H}^t(\,\mathbf{H}\,\mathbf{P}_\mathbf{0}\,\mathbf{H}^t + \mathbf{R}_\mathbf{v}\,)^{-1}\,(\mathbf{g} - \mathbf{H}\,\mathbf{f}_0) \\ &= \mathbf{f}_0 + \mathbf{A}\,(\mathbf{g} - \mathbf{H}\,\mathbf{f}_0),\end{aligned} \qquad (4)$$

It is still assumed that the sub-imagers are well focused, and that the point spread function of the optical system is small compared to the area of individual pixels. In general the **H** matrix will represent the integrating effect of the pixel sensors for a given lens magnification. A number of adjacent pixels in **f** will be added with approximately uniform weighting to form each observed pixel in **g**. If the response of the lens system is shift invariant, the resulting null space makes this reconstruction problem ill conditioned and will limit performance in the presence of noise or limited quantization.

If Gaussian statistics are assumed, $\mathbf{R}_\mathbf{f}$, the expected covariance of the image error ($\mathbf{f}-\hat{\mathbf{f}}$), can be written as shown in Equation (5). Here $\mathbf{R}_{\mathbf{vM}}$ is the actual covariance matrix for the noise vector **v**, and $\mathbf{P}_{\mathbf{0M}}$ is the actual covariance matrix for the initial estimate of the image data. This can be used to compare the expected performance of different reconstruction methods.

$$\mathbf{R}_\mathbf{f} = (\mathrm{I} - \mathbf{A}\mathbf{H})\,\mathbf{P}_{\mathbf{0M}}(\mathrm{I} - \mathbf{K}\,\mathbf{H})^{\,t} + \mathbf{A}\,\mathbf{R}_{\mathbf{vM}}\,\mathbf{A}^{\,t} \qquad (5)$$

IV. Benefits of Lens Diversity

The performance of the image estimator of Equation (4) can be improved by taking additional measurements. A second set of measurements for the same **f** taken with the same type of sub-imagers will help by averaging noise, but will not add much new information. However, if a different imaging geometry is used for a second set of measurements, there will be much greater improvement in the reconstructed image. Using several different sub-imager arrays with different imaging system characteristics, as shown in Equation (6), can improve performance by reducing the null space associated with a single lens type.

$$g = \begin{bmatrix} g_I \\ g_{II} \\ g_{III} \end{bmatrix} = \begin{bmatrix} H_I \\ H_{II} \\ H_{III} \end{bmatrix} f + v \qquad (6)$$

Using a simple example, Figure 2 demonstrates the improvement in the average expected squared error for reconstructions of image tiles using lens diversity. Sub-imager array type I uses a 3x3 array of sub-imagers in which each sub-imager has a 5x5 pixel sensor and the magnification is 1/3. Sub-imager array type II uses a 5x5 array of sub-imagers in which each sub-imager has a 3x3 pixel sensor and the magnification is 1/5. Sub-imager array type III uses a 4x4 array of sub-imagers in which each sub-imager has a 4x4 pixel sensor and the magnification is 1/4.

Figure 2 plots the average expected squared error for a reconstructed 19x19 pixel image tile as a function of the actual variance of the added noise. Estimators used expected noise values of 10^{-6}, 10^{-4}, 10^{-2}, and 1. It is assumed **f** is selected from the set of all possible images with pixel values between 0 and 255. The top curve shows the error from two sets of sub-imagers, both of type I. The curves for all four estimators are almost identical and can not be distinguished for actual noise variances less than 1. There is almost no improvement in the expected error as the noise variance is reduced. The next two solid curves show the error for a combination of two sets of sub-imagers, one of type I and one of type II. For actual noise variances below 0.1, the two-lens system outperforms a single lens system using the same number of observations. In both cases, the null space of H prevents improvement in the expected error for very low values of noise variance.

The lower three solid-line plots show the expected error for a combination of all three types of lens systems. The expected error is reduced as the actual noise variance is reduced until it falls below the expected variance.

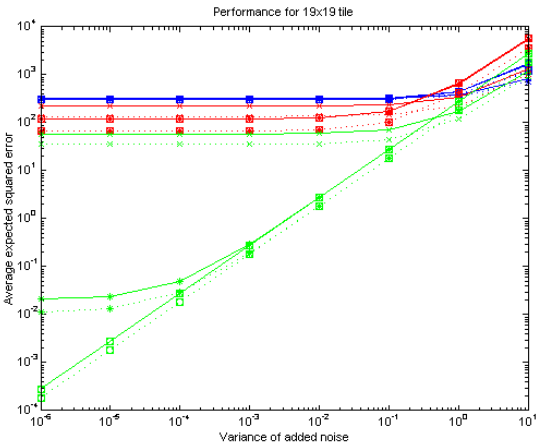

Figure 2: Expected squared error as a function of measurement noise variance for a single image tile for different lens types and expected noise variance.

When two lens types are used, the expected error in the estimates of the image pixels is not uniformly distributed over the 19x19 pixel tile, and the pixels at the edge of the tile have the lowest confidence. This is demonstrated for the low noise case in Figure 3 in a two dimensional self normalized display of expected error for an image tile. The maximum value associated with a white pixel appears above each display. For the type I lens only, shown on the left, there is no reduction in the average error when the border pixels are removed. For the combination shown in the center, the highest errors are on the edges, and trimming the borders of the reconstructed tile reduces the average error. The relative improvement from trimming the borders is less when three different magnifications are used, but the average expected error is much lower in both cases. The dotted curves in Figure 2 show the average expected error of the trimmed tiles for the various estimators.

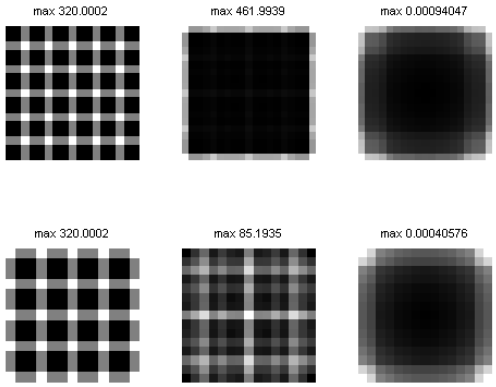

Figure 3: Spatial distribution of expected squared error within the tile for type I lens (left), combined I and II (center) and all three types (right). The top row shows results for the full tile and the bottom row has the border pixels trimmed away.

When a larger sensor array is used, but all other factors are constant, the reconstructed image will have a lower expected error, but the computation defined by Equation (4) will increase significantly. In Figure 4 the expected error for a 35x35 tile is plotted in the same manner as shown in Figure 2. The relative improvement using the larger tile size is more apparent for the systems where the null space of H limits performance.

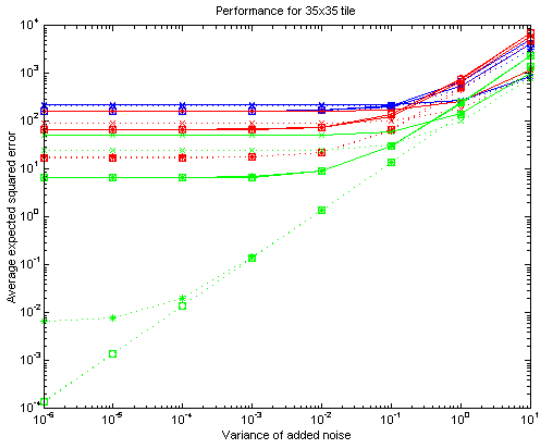

Figure 4: Results shown as in Figure 2 for a 35x35 tile instead of 19x19.

When the magnification is reduced, the noise variance must also be reduced to achieve the same expected error

performance. Figure 5 compares the performance of the three-lens systems shown in Figure 2 for a 19x19 tile and Figure 4 for a 35x35 tile with the performance of a 41x41 tile reconstructed from images with magnifications of 1/5, 1/6, and 1/7. The parallel straight diagonal lines show the error for the estimator with the lowest expected noise variance. The lowest line, for the 35x35 tile, shows the small improvement when the tile size is increased compared to the line for the 19x19 tile just above it. The highest of the three lines is the 41x41 tile reconstructed from lower resolution images. For this case, the noise variance must be reduced by approximately a factor of 10 to maintain the same expected error.

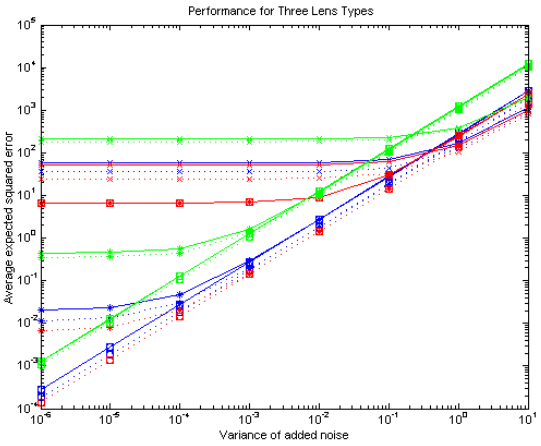

Figure 5: Comparison of the three-lens systems from Figures 2 and 4 with a 41x41 tile using magnifications of 1/5, 1/6, and 1/7.

V. Computational Issues

Results from simulations are consistent with the expected results shown in Figures 2 through 5. The bridge image in Figure 6 from the USC database was divided into small overlapping tiles for reconstruction. Several with different average values, contrast, and content are shown. Simulated measurements used added measurement noise or quantized measurement values. The pixels of the image source are not uncorrelated as modeled, and the actual average performance in simulation is somewhat better than the predictions.

Figure 7 shows plots of the average squared error for simulated reconstructions of multiple image tiles as a function of measurement noise. In Figure 7(a) the variance of the added noise is varied but there is no quantization noise. In Figure 7(b) there is no added measurement noise but the number of bits used to represent the measured values in **g** is varied. The quantization effects can be interpreted as adding noise with a variance determined by the values of the least significant bit of the quantized measurements, and this interpretation is consistent with the results in Figure 7(a).

Figure 6: Original 512x512 image and six tiles

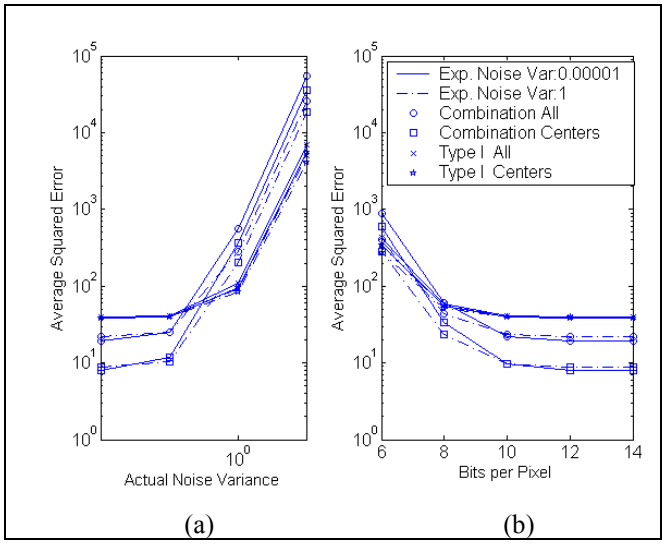

Figure 7: Average squared error as a function of measurement noise variance (a) and number of bits per measurement (b).

There is a small overlap in the fields of view for pixels used to compute the values of adjacent individual tiles, so some further improvements could be attained if larger tiles were used in the reconstruction. This was demonstrated in Figure 5. However, using larger tiles would significantly increase the computational requirements. The same improvements can also be achieved by combining results from smaller tiles computed independently and in parallel. The low rank corrections using the ABCD Lemma in Equation (7) could combine the information from adjacent tiles. The H matrix for a combination of several tiles would have low rank connections between the smaller H matrices corresponding to the overlap in fields of view of the border pixels of the smaller tiles.

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1} \qquad (7)$$

Note that the general statement of the lemma does not refer to the matrix A used in Equations 4 and 5.

VI. Conclusions

Arrays of sub-imagers which produce many low resolution images with large amounts of overlap in field of view can be used to create an estimate of a higher resolution image. This is a well known theoretical problem in super resolution, but in practical implementation, the results are often not as good as expected.

This paper shows that using lens diversity in the form of several different lens types or a lens with a space variant magnification improves the quality of the result and allows the computation to be partitioned for parallel processing. In the example used to demonstrate these results, the second added lens system had a lower resolution than the first, but it reduced the null space of the reconstruction. The results of the two lens system had a lower error than the results obtained from a system which added a second set of the same higher resolution lens arrays. Adding a third lens system with a different magnification eliminated the null space of H and the expected error performance was limited only by the variance of the measurement noise. As expected, lowering the resolution of the sub-imagers requires a reduction in the noise variance to keep the same expected error performance.

Using lens diversity allows small tiles of a larger image to be computed with little penalty in increased expected error. Although increasing the size of the reconstructed image reduces the expected squared error, it also significantly increases the computational cost. The same improvement in expected error can be obtained with an update based on the low rank connection between adjacent tiles.

Both theoretical computation of average expected squared errors and simulations of reconstructions in the presence of added measurement noise and limited measurement precision demonstrate the advantages of lens diversity for image reconstruction. This can be extended to derive strategies for adding new measurements which will result in the most improvement. The effects of inaccuracies in H due to misalignment or modeling assumptions or signal dependent noise are left to future studies.

VII. References

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

[2] L. G. Brown, "A Survey of Image Restoration Techniques," *ACM Computing Surveys*, vol. 24, no.4, pp325-376, December 1992.

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

[4] M. P. Christensen, M.W. Haney, D. Rajan. S.L. Wood, S. Douglas, "PANOPTES: A Thin Agile Multiresolution Imaging Sensor", to appear in *Proc GOMACTech2005*, Las Vegas, NV, April, 2005.

[5] D. Daly, *Microlens Arrays*, Taylor and Francis, N.Y.. 2001.

[6] M.W. Haney, M.P. Christensen, D. Rajan, S. Douglas, S.L. Wood, "Adaptive Flat Micro-mirror Array-based Computational Imaging Architecture", to appear in *Proc. OSA COSI*, June 2005.

[7] Y. Kitamura, R. Shogenji, K. Yamada, S. Miyatake, M. Miyamoto, T. Morimoto, Y. Masaki, N. Kondou, D. Miyazaki, J. Tanida, Y. Ichioka, "Reconstruction of a High-resolution Image on a Compound-eye Image-capturing System", *Applied Optics*, Vol 43, No. 8, 10 March 2004.

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

[9] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, Y. Ichioka, "Thin Observation Module by Bound Optics (TOMBO): Concept and experimental verification," *Appl. Optics-IP*, vol. 40, no. 11, p. 1806, Apr. 2001.

[10] S. Wood, D. Rajan, M. Christensen, S. Douglas, B. Smithson, "Resolution Improvement for Compound Eye Images through Lens Diversity," *Proc. 11th DSP Workshop*, Taos, NM, August 1-4, 2004.

[11] S. Wood, B. Smithson, D. Rajan, M. Christensen, "performance of a MVE Algorithm for Compound Eye Image Reconstruction using Lens Diversity," *Proc. ICASSP2005*, Philadelphia, Pa, 20-23March2005.

Main Menu