

PERFORMANCE OF A MVE ALGORITHM FOR COMPOUND EYE IMAGE RECONSTRUCTION USING LENS DIVERSITY

Sally L. Wood¹ Bonnie J. Smithson¹ Dinesh Rajan² Marc P. Christensen²

¹Santa Clara University
Santa Clara, California, 95053 USA

²Southern Methodist University
Dallas, Texas, 75275 USA

ABSTRACT

Reconstruction algorithms to compute a single improved resolution image from multiple lower resolution images have application in the design of cameras with flat form factors. The accuracy of these reconstructions will depend on measurement noise, measurement quantization, the structure of the image acquisition system, and the accuracy of the image acquisition model. This paper compares the expected and simulated performance for reconstructions from multiple lower resolution images. The analysis shows that designs using lenses with different imaging characteristics significantly improve the theoretical performance results. In addition, lens diversity allows the reconstruction problem to be naturally partitioned into a set of loosely coupled smaller reconstructions that are computationally more manageable.

1. INTRODUCTION

A typical optical imaging device uses a single lens to focus the desired image onto a sensor array and then digitizes the output of each sensor. The bulk of the cost associated with this approach is usually in the optical elements. The length requirements of the optical column preclude the use of a flat form factor for the device, but flat cameras have many potential applications ranging from conformal sensing skins on surveillance aircraft to head-mounted camera patches for firemen and rescue workers.

In contrast to the traditional imaging approach, in radio astronomy or medical imaging [1,6] the image field is sampled with lower resolution elements, and image reconstruction methods are used to compute better quality images. Recent advances in optical imaging sensors have created an opportunity to bridge these two imaging approaches to design devices with a flat form factor [5]. For example, the TOMBO imaging device [7] 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 [4,8]. This type of reconstruction task uses methods similar to those of computed tomography in medical imaging [6] and image registration [e.g. 2].

Reconstruction of high resolution images from multiple overlapping low resolution images is a well-known theoretical super-resolution problem [3]. 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. Theoretically, small translations are desired for high resolution, but small translations also result in a small amount of new information from each additional image and require a higher bit precision in the measurements to represent small but significant differences.

The objective of the approach described here 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. In this paper, we explore the use of lens diversity and matched partitioning using the minimum variance estimator. Performance of the reconstruction algorithm is improved and made more robust by adding observations taken from sensor/lens systems with different magnifications. 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. The technique allows for parallel computation and merging of multiple reconstruction results. Both expected error computations and simulated reconstruction examples demonstrate the advantages of our approach.

2. MODEL OF THE IMAGING SYSTEM

In a traditional single lens high resolution imager, the observed image \mathbf{g} is related to the desired image \mathbf{f} by an observation matrix \mathbf{H} as shown in Equation (1).

$$\mathbf{g} = \mathbf{H} \mathbf{f} + \mathbf{v}. \quad (1)$$

The two dimensional images are stored in the column vectors \mathbf{f} and \mathbf{g} in row order, and vector \mathbf{v} represents uncorrelated measurement noise. The observations in \mathbf{g} are viewed as our estimate of the desired image \mathbf{f} . Sometimes image restoration may also be used when statistics of the noise and image source class are known.

The \mathbf{H} matrix represents the model of the imaging system. Typically the system is in focus, and the point spread function of the imaging system is assumed to be small compared to pixel size. The resolution of \mathbf{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

The lens just described requires a ‘‘cubic’’ or large form factor. A flat camera form factor would have to use a smaller lens positioned closer to a smaller sensor array. The change in magnification factor would reduce the resolution of \mathbf{g} . Equation (1) would still be the model, but the size of \mathbf{g} would be much smaller than the traditional system and \mathbf{H} would be modified for the characteristics of the small lens.

An array of small imaging systems, called sub-imagers, could provide information needed to improve the resolution of \mathbf{g} . Each sub-imager would have a micro lens 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 \mathbf{f} will have $N = K_f * L_f$ elements.

Let \mathbf{g}_{00} be the J element pixel array associated with a single sub-imager, where J is much smaller than N . A second identical sub-imager, \mathbf{g}_{01} , is 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 \mathbf{Z} with elements $z(k, l) = \delta(k + 1 - l)$. Similarly, let \mathbf{g}_{10} be another identical sub-imager with a vertically shifted field of view. The model for the shifted images is given by Equation (2). It is assumed that \mathbf{f} spans the field of view for all sub-imagers.

$$\begin{aligned} \mathbf{g}_{00} &= \mathbf{H}_s \mathbf{f} + \mathbf{v}_{00} = \mathbf{H}_{00} \mathbf{f} + \mathbf{v}_{00}. \\ \mathbf{g}_{01} &= \mathbf{H}_s \mathbf{Z}^X \mathbf{f} + \mathbf{v}_{01} = \mathbf{H}_{01} \mathbf{f} + \mathbf{v}_{01}. \\ \mathbf{g}_{10} &= \mathbf{H}_s \mathbf{Z}^Y \mathbf{f} + \mathbf{v}_{10} = \mathbf{H}_{10} \mathbf{f} + \mathbf{v}_{10}. \end{aligned} \quad (2)$$

When multiple small images formed with multiple small lenses are used to create an estimate of the desired image, it can be considered a problem of image reconstruction, because new observations can be used to add new information about \mathbf{f} . A $K \times L$ 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]. Equation (3) represents the combined observations from the array of sub-imager sensors in the same form as Equation (1). However, in this case, the array of low-resolution images contained in \mathbf{g} does not represent the final image we wish to view. These measurements must be used to compute a higher resolution estimate of \mathbf{f} .

$$\begin{aligned} \mathbf{g} &= [\mathbf{g}_{00}^t \ \mathbf{g}_{01}^t \ \dots \ \mathbf{g}_{0L-1}^t \ \mathbf{g}_{10}^t \ \dots \ \mathbf{g}_{K-1L-1}^t]^t \\ &= [\mathbf{H}_{00}^t \ \mathbf{H}_{01}^t \ \dots \ \mathbf{H}_{K-1L-1}^t]^t \mathbf{f} + \mathbf{v} = \mathbf{H} \mathbf{f} + \mathbf{v} \end{aligned} \quad (3)$$

A linear minimum variance of error estimate for \mathbf{f} can be computed using Equation (4), where \mathbf{f}_0 is the expected average value for the image, \mathbf{R}_v is the covariance matrix for the noise vector \mathbf{v} , and \mathbf{P}_0 is the initial image covariance matrix.

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

In general the lens characteristics represented by \mathbf{H} will add a number of adjacent pixels in the original image with approximately uniform weighting to form each observed pixel in \mathbf{g} . The resulting null space makes this reconstruction problem ill conditioned and will limit performance.

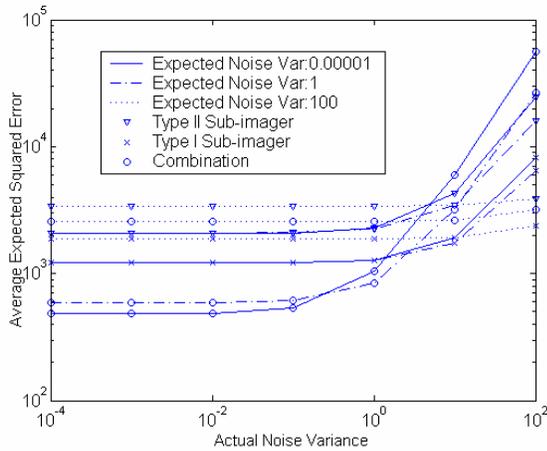
3. DIVERSITY IN IMAGE MAGNIFICATION

The performance of the image pixel estimator of Equation (4) can be improved by taking additional measurements. A second set of measurements for the same \mathbf{f} taken with the same type of sub-imagers will help in 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 two different sub-imager arrays with different imaging system characteristics as shown in Equation (5) can improve performance by reducing this null space.

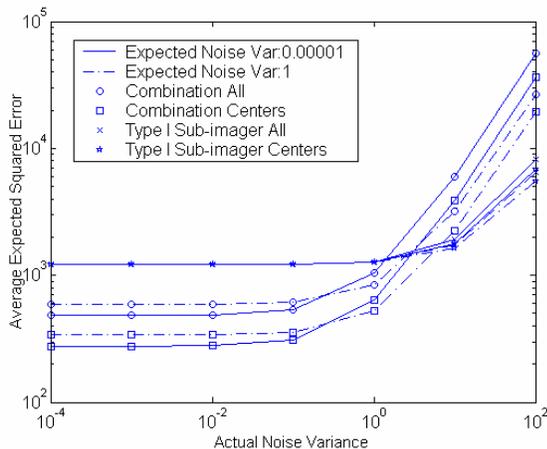
$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_I \\ \mathbf{g}_{II} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_I \\ \mathbf{H}_{II} \end{bmatrix} \mathbf{f} + \mathbf{v} \quad (5)$$

Figure 1 demonstrates the improvement in the average expected squared error for reconstructions of image tiles using lens diversity. Sub-imager array type I uses a 3×3 array of sub imagers in which each sub-imager has a 5×5 pixel sensor and each pixel is the average of a 3×3 pixel

region of the original image. The lower resolution type II sub-imager array uses a 5x5 array of sub imagers in which each sub-imager has a 3x3 pixel sensor and each pixel is the average of a 5x5 pixel region of the original image.



(a)



(b)

Figure 1: Average expected squared error as a function of measurement noise variance for several reconstruction approaches for a single image tile. Both full and trimmed reconstructions are shown.

Figure 1(a) plots the average expected squared error for a reconstructed 19x19 pixel image tile as a function of the variance of the added noise. It is assumed \mathbf{f} is selected from the set of all possible images with pixel values between 0 and 255. For each imaging geometry three estimators using different values for estimated noise variance were tested. When the expected noise variance is 100, there is a high average expected error in $\hat{\mathbf{f}}$ even when the actual noise variance is low. When the expected noise variance is close to the actual value, the average expected squared error is reduced by 60% in the low noise

case when both lens types are used instead of the single type I lens. The same number of observations is used in both cases. There is a corresponding reduction of 33% when the noise variance is 1. Addition of a third lens type further reduces error for the low noise case.

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. In Figure 1(b) some results from Figure 1(a) are compared to results trimmed to just the 15x15 center of the 19x19 tile. There is noticeable improvement in the average expected error for combined lens systems when only the 15x15 center of the reconstruction is used. For the two lens system, when the expected noise variance is close to the actual value, the average squared error is reduced by 45% in the low noise case and by 37% when the noise variance is 1. This shows that independent computations for small overlapping tiles can be used for a good first approximation when there is lens diversity, and this will greatly reduce computational complexity. The results for a single lens system do not show any significant advantage in trimming the edge pixels.

Results consistent with those of Figure 1 are obtained from simulations using added measurement noise or quantized measurement values. The bridge image in Figure 2 from the USC data-base was divided into small overlapping tiles for reconstruction.

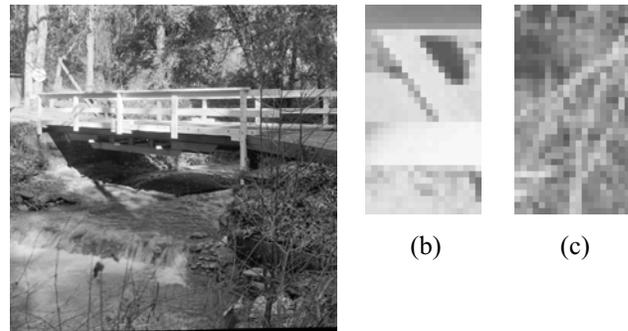


Figure 2: Original 512x512 image in (a) is reconstructed in tiles of 19x19, 19x34, or 34x34 pixels. Two 19x34 tiles are shown in (b) and (c).

Figure 3 shows plots of the average squared error for simulated reconstructions of multiple image tiles from Figure 2 as a function of the variance of the added noise and the number of bits used to represent the measured values in \mathbf{g} . 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. The performance for individual tiles varied depending on the actual content. Reconstructions of random patterns

matching the assumptions used for Figure 1 yield the predicted results.

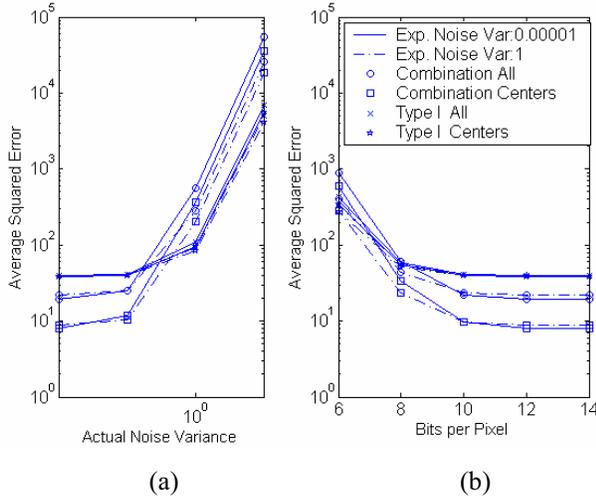


Figure 3: Average squared error as a function of measurement noise variance (a) and number of bits per measurement (b) for image tiled from Figure 2.

4. COMPUTATION OF LOW RANK UPDATES

Figure 4 shows that using twice as many measurements to compute estimates of image tiles that are twice as large further reduces the expected average squared error when two types of lens systems are used. The estimates made for 30x15 centers of 34x19 tiles are compared to the results shown in Figure 1(b). When three types are used the initial local estimate is better and thus the relative improvement from the larger local area is less noticeable.

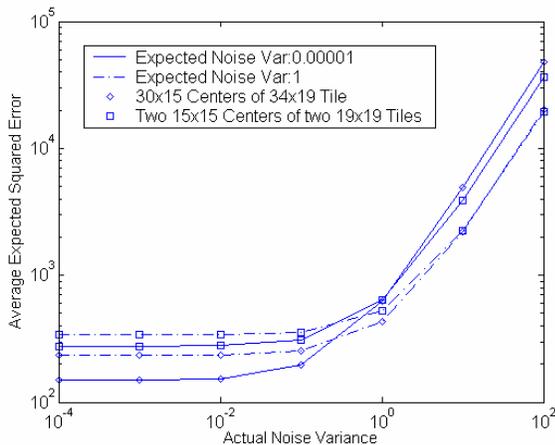


Figure 4: Average expected squared error for reconstruction of 30x15 pixel tiles compared to error for 15x15 pixel tiles.

This improvement can be accomplished by using Equation (4) directly with twice as many measurements, or by first computing the estimates for the individual

19x19 tiles and using low rank updates based on Equation (6) to update each small tile based on the other. Here A represents two uncoupled tiles. When the expected noise variance is close to the actual value, the average squared error is reduced by 45% in the low noise case and by 20% when the noise variance is 1.

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} \quad (6)$$

5. CONCLUSIONS

Reconstructing a high resolution image from multiple lower resolution images is a well known theoretical problem in super resolution. In practical implementation, the results are often poorer than expected. This paper shows that improvement in the quality of the results and reduction of the computational load can be achieved when a combination of different lens systems is used to acquire the low resolution images. Although the second added lens system has a lower resolution than the first, it reduces the null space of the reconstruction and allows the problem to be partitioned for more manageable local computations. Adding a third lens system can eliminate the null space.

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. Future studies will consider the effect of inaccuracies in H due to misalignment or modeling assumptions.

6. REFERENCES

- [1] R. N. Bracewell, *Two-Dimensional Imaging*, Prentice Hall, Englewood Cliffs, N.J., 1995.
- [2] L. G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, no.4, pp325-376, Dec. 1992.
- [3] S. Chaudhuri, ed. *Super-Resolution Imaging*, Kluwer Academic Publishers, Boston MA, 2001.
- [4] M. Christensen, M. Haney, D. Rajan, S. Wood, S. Douglas, "PANOPTES: A thin agile multi-resolution imaging sensor," invited GOMACTech 2005.
- [5] D. Daly, *MicroLens Arrays*, Taylor and Francis, N.Y., 2001.
- [6] A. Macovski, *Medical Imaging Systems*, Prentice Hall, Englewood Cliffs, N.J., 1983.
- [7] 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.
- [8] 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.