IMPROVED MULTIPLEXED IMAGE RECONSTRUCTION PERFORMANCE THROUGH OPTICAL SYSTEM DIVERSITY DESIGN

Sally L. Wood[1], *Hsueh-Ban Lan*[1],*Dinesh Rajan*[2], *Marc P. Christensen*[2]

[1]Department of Electrical Engineering, Santa Clara University
500 El Camino Real, Santa Clara, California, 95053
swood@scu.edu, benh_6_lan@yahoo.com
[2]Electrical Engineering Department, Southern Methodist University
6251 Airline Road, Dallas, Texas, 75275
rajand@engr.smu.edu, mpc@engr.smu.edu

ABSTRACT

Multiplexed image reconstruction, estimating high resolution images from multiple low resolution images with highly overlapped fields of view, is improved when the magnification of the imagers is diverse. No assumptions of shift invariance or Toeplitz structure are required for computational manageability because localized reconstruction is possible and sensitivity to boundary conditions is reduced. Such multiplexed diverse image sensors have applications in flat sensor systems for surveillance and pervasive personal imaging.

Index Terms— Multiplexed imagers, image reconstruction, computational imaging, flat camera.

1. INTRODUCTION

High resolution image reconstruction from multiple low resolution images is important for two different application areas of current interest: i) A low-cost low resolution camera that could produce images equal in quality to those produced by a more expensive higher resolution camera would have numerous commercial applications in imaging and video [1–3]. ii) A computational imaging system with a flat form factor may require using smaller optical elements positioned closer to the image detectors [4–9]. This creates a physically smaller image on the detector array for the same field of view (FOV). Proportionately reducing the size of the detector elements to maintain high resolution is not feasible because it would decrease the SNR and cause manufacturing difficulties.

In both applications, the higher resolution image would be created computationally from multiple low resolution images with sub-pixel translations. This super-resolution problem is well known, ill-posed, and computationally intensive [10–13]. The reconstruction methods are similar to well-known algorithms used in radio astronomy (*e.g.* [14]) and medical imaging (*e.g.* [15]). The performance of an optimal reconstruction method will depend on the noise level in the low resolution images and the degree to which the information in the desired higher resolution image is captured by the sequence of low resolution images.

This paper will address the flat camera reconstruction problem where each low resolution image is captured by an individual small imaging device called a sub-imager (SI). Reducing the noise level of the SIs and providing magnification diversity can extend the limits in [13], which were based on a single magnification. A single magnification was also used in the TOMBO system [5] in which a conventional 240×240 pixel sensor array was divided into a 6×6 array of 40×40 pixel sub arrays, each of which had a $500\mu m$ aperture micro-lens with f=1.3 mm. The object distance was fixed at 26 cm, and a resolution improvement by a factor of 4 was reported. Reconstruction results have also been demonstrated using a single low resolution camera to provide an image sequence [1–3]. However, performance analysis for this application could depend on differences between the low-cost and high-cost cameras in the optics quality and the detector efficiency as well as the camera resolution difference.

Previous results have demonstrated the performance improvement due to optical system diversity for PANOPTES, a new flat-form factor computational imaging architecture [6, 8]. In this paper these results are extended to non-integer magnification ratios, reduced diversity systems, and space variant magnifications. Diversity in the effective magnification of an imaging system relative to the sampled pixel array can be created in several ways. A basic micro-lens/micro-mirror array designed to meet the flat camera thickness specifications could be augmented with additional imaging resources that would project a larger field of view onto a pixel than the basic array. Alternatively, added optical elements of the same type as the basic array could project the image source onto physically larger pixels which could also have lower noise levels. A third possibility would use arrays of optical elements with space variant magnification.

The paper is organized as follows. In Section 2 we introduce the basic model used in the analysis. Diversity aspects of the proposed imaging system are discussed in Section 3 and the performance analysis is presented in Section 4.

2. MATHEMATICAL MODEL

A continuous image source $f_C(x, y)$, bandlimited by the point spread function of the optical system, is projected onto an array of rectangular detectors. The point spread function is assumed to be small compared to the detector size. The convolution of $f_C(x, y)$ with the rectangular shape function of the detector produces $g_C(x, y)$, which is sampled at the detector spacing. For a detector of width a and height b, the transform of the shape function to spatial frequency (u, v) is $(\sin(\pi ua)/(\pi u))(\sin(\pi vb)/(\pi v))$, which will produce periodic nulls in the frequency response. In the absence of known boundary conditions, it is not feasible to restore the zeroed frequencies by combining images acquired with sub-pixel translations.

For a one-dimensional image model, it is possible to reconstruct a high resolution image at all frequencies if images are acquired

from two different systems using detector window widths that are coprime in terms of the coordinates of $f_C(x, y)$. For a two-dimensional model, three different window widths are needed [8]. Selection of the window widths for optimal performance must consider both the positions of the zeros in the frequency response and the need to lower the measurement noise as the size of the window is increased to maintain the same expected error performance.

Figure 1 shows the frequency response for two systems with three different magnifications. Let q represent the window width in units of the desired pixel resolution of the reconstructed image. In Figure 1a, the values of q are 3, 4, and 5. The zeros are well distributed along the normalized frequency axis, but the width of the largest window is 67% larger than the base window width. In contrast, in Figure 1b, the values of q are 3, 3.2, and 3.4. Consequently, the zeros are closely distributed in the normalized frequency axis, but now the width of the largest window is only 13% larger than the base window width. The desired discrete source image is modeled

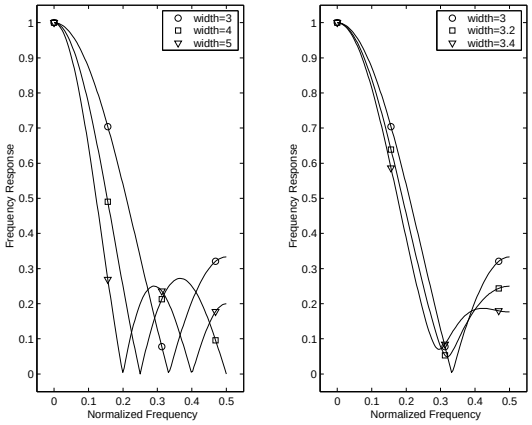

Fig. 1. Frequency response of two systems with diverse magnification

as a pixel array at the desired resolution. Let F represent the N_y row and N_x column image source with $N = N_y N_x$ source pixels. This source image array is stored by rows in an $N \times 1$ column vector, f. Similarly, the $M_y \times M_x$ sensor image, G is stored by rows in the $M \times 1$ column vector, g.

An observation partial matrix H models the behavior of the optical elements, the integration of the detector array sensors, and the physical geometry of the sensor array. A linear mathematical model of the measurement system is considered as given by,

$$g = Hf + v, \tag{1}$$

where, noise vector v includes uncorrelated random noise from a variety of sources. When the measurement noise and the class of possible image sources are modeled as Gaussian random variables, the minimum variance estimate (MVE), $\hat{f}$, can be computed as [16],

$$\hat{f} = f_0 + K(g - Hf_0) \tag{2}$$

The matrix K is defined using estimates of the noise covariance, $\hat{R}_v = \mathbb{E}[vv^T]$, and the covariances of the image class, $\hat{P}_0 = \mathbb{E}[(f - f_0)(f - f_0)^T]$ as,

$$K = \hat{P}_0 H^T (H\hat{P}_0 H^T + \hat{R}_v)^{-1} \tag{3}$$

This derivation of K assumes that the noise has a zero mean value and that the image and noise are uncorrelated, *i.e.*, $\mathbb{E}[fv^T] = 0$.

The covariance of the estimate error $\xi = \mathbb{E}[\tilde{f}\tilde{f}^T]$ where $\tilde{f} = f - \hat{f}$ can be used to find the expected squared error for an imaging system modeled by H. The variance of each individual pixel in $\hat{f}$ can be computed using (4) for any K. Here actual noise covariance R_v and the actual image class covariance P_0 may be different from the estimated values. The estimation error is given by,

$$\xi = \mathbb{E}(\tilde{f}\tilde{f}^T) = (I - KH)P_0(I - KH)^T + KR_vK^T \tag{4}$$

The diagonal elements of the matrix ξ contain the expected variance for each element of $\tilde{f}$.

Both analytical analysis and simulation results show that when H models a low resolution sensor array, the spatial frequencies that average to zero across the sensor can never be recovered. Hence, continued reduction of the measurement noise does not result in comparable reductions in the average expected squared error [6, 8].

3. ARRAYS OF DIVERSE SUB-IMAGERS

Since no assumption of shift invariance was used, the mathematical formulation in the previous section applies equally well to the case where H represents an array of identical SIs or an array of diverse SIs. In both cases a high resolution reconstruction can be computed from an array of SIs with subpixel relative translations of sensor arrays. The expected error of the reconstruction can be reduced by using diverse SIs, and different types of diverse sub-imager arrays can be compared.

The mathematical representation of a single SI in the previous section can be extended to represent an array of SIs. Let $H_{0,0} = H$ represent a single SI in which the detector FOVs are not overlapping and each detector averages a $q \times q$ pixel area of the desired image. A $q \times q$ array of sub-imagers with FOVs offset from each other by one desired source pixel width will have a combined FOV that is extended by $q - 1$ pixels horizontally and vertically with $N_{xE} = N_x + q - 1$ and $N_{yE} = N_y + q - 1$. The shift matrix Z_l is an $N \times N_E$ matrix in which $Z_l(i, j) = \delta(i - (j - l))$. If the H matrix of a SI is post-multiplied by Z_l, the FOV of the SI will effectively shift to the right by l pixel position in the f vector and will place the SI's FOV correctly in the extended composite FOV of all SIs. Row shifts of k rows will require a multiplication by $Z_{kN_{xE}}$ for the two-dimensional image data stored in vector f.

The vector $g_{k,l}$ represents the detector image from the $(k, l)^{th}$ SI. Using the measurement model of (1), individual equations can be written for each SI as,

$$g_{k,l} = HZ_{kN_{xE}+l}f + v_{k,l} = H_{k,l}f + v_{k,l} \tag{5}$$

The total number of observed pixels from the sub-arrays is the same as the number of pixels in a traditional high resolution imager. However, the total FOV for the array of SIs is extended because no assumptions are made about boundary conditions.

The column vectors of image outputs from all the SIs in (5) can be combined into a single array of data as follows:

$$\begin{aligned} g &= H_q f + v \\ g^T &= [g_{0,0}^T,\ g_{0,1}^T,\ \cdots,\ g_{q-1,q-1}^T]^T \\ H_q^T &= [H_{0,0}^T,\ H_{0,1}^T,\ \cdots,\ H_{q-1,q-1}^T]^T \\ v^T &= [v_{0,0}^T,\ v_{0,1}^T,\ \cdots,\ v_{q-1,q-1}^T]^T \end{aligned} \tag{6}$$

The subscript q on the combined H matrix indicates a $q \times q$ array of SI with linear resolution reduced by a factor of q.

The expected squared error for the MVE reconstruction can be estimated using the singular value decomposition or spectral representation [17]. Let $H_q = USV^T$, where U and V are $N \times N$ and $N_E \times N_E$ unitary matrices and S is an $N \times N_E$ matrix with singular values along the diagonal. Assume that $R_v = \sigma^2 I_N$ and $P_0 = p_0 I_{N_E}$. Then $H_q P_0 H_q^T = p_0 U S S^T U^T$ and, using (4) $\xi = V \Lambda V^T$ where the elements of the $N_E \times N_E$ diagonal matrix Λ are given by (7). The elements of Λ determine the average expected squared error when $\hat{f}$ is computed using the MVE with the detector images in (6) and are given by,

$$\lambda_i = \begin{cases} \frac{p_0 \sigma^2}{p_0 s_i^2 + \sigma^2} & \text{for } 1 \le i \le N \\ p_0 & \text{for } N+1 \le i \le N_E \end{cases} \tag{7}$$

The performance improvements due to magnification diversity can be analyzed using these results for a SI array. An H_q matrix is generated for each magnification, and the observed images for arrays of SI's at the three magnification levels are treated as a single set of observations as follows,

$$g = \begin{bmatrix} g_{q1} \\ g_{q2} \\ g_{q3} \end{bmatrix} = \begin{bmatrix} H_{q1} \\ H_{q2} \\ H_{q3} \end{bmatrix} f + v = Hf + v \tag{8}$$

The performance of this diverse system can be analyzed using,

$$\hat{f} = (H^T R_{v_e}^{-1} H + P_{0_e}^{-1})^{-1} (H^T R_{v_e}^{-1} g + P_{0_e}^{-1} f_0) \tag{9}$$

If $H^T = [H_{q1}^T \, H_{q2}^T \, H_{q3}^T]$ then $H^T R_{v_e}^{-1} H$ can be written as

$$H^T R_{v_e}^{-1} H = H_{q1}^T R_{v_{e_{q1}}}^{-1} H_{q1} + H_{q2}^T R_{v_{e_{q2}}}^{-1} H_{q2} + H_{q3}^T R_{v_{e_{q3}}}^{-1} H_{q3} \tag{10}$$

If each H_q represents a large shift invariant convolution, then the circulant approximation to Toeplitz structures can be applied [17, 19]. In that case the eigenvectors of all three $H_{qi}^T H_{qi}$ matrices will be the same DFT vectors. This allows the matrices in (10) to be added by simply adding the eigenvalues of each of the $H_{qi}^T H_{qi}$ matrices scaled by the appropriate inverse noise variance. This matches the intuitive concept of adding new types of SIs to capture all spatial frequencies and demonstrates the value of diversity.

However, without any shift invariant assumptions, magnification diversity still allows small over-determined tiles of an image to be reconstructed independently. For the small tiles the circulant approximation is not appropriate, and more exact methods are used to compute K [8, 18]. The expected error computed for small tiles in the next section is based on computation of slightly overlapping small image tiles. The overlapping edges of the reconstructed tiles with high expected error are discarded [8, 18].

4. EXPECTED PERFORMANCE

Figure 2 compares single and diverse magnification systems in terms of the expected mean squared error as a function of actual measurement noise. A maximum image value of 255 is assumed. Results for four systems are shown, and for each system four estimators are shown corresponding to four different values of expected noise variance indicated in the legend. The best performance is shown by the solid line plots for a system with q of 3, 4, and 5. The worst performance, shown by the dotted line plots, corresponds to a system using only $q = 3$. At this scale, results for all four estimators with $q = 3$ are superimposed across the top of the plot. The same number of observations was taken for both the $q = 3$ system and the $q = 3, 4, 5$ system. It is clear that without magnification diversity, the reconstruction algorithms can not reduce the average expected squared error as the measurement noise is reduced. The other two systems shown use all three magnification values, but each only uses 33% of the imagers for $q = 5$. This increases the error by approximately a factor of 2 at low noise levels, and it is still superior to the single magnification system.

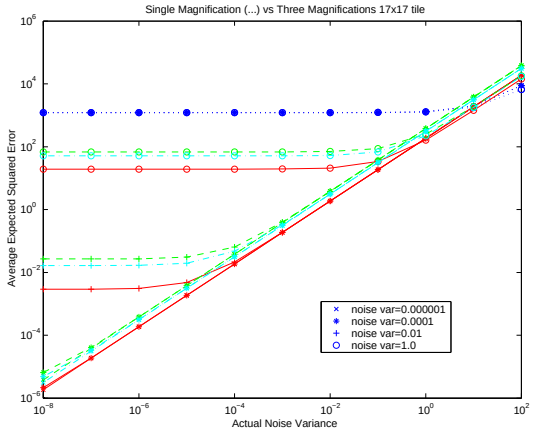

Fig. 2. Effect of reducing the density of the elements with higher values of q

Figure 3 shows results for systems with less magnification diversity. For reference, the solid and dotted lines correspond to the same systems described in Figure 2. The three additional systems have the following sets of q values: {3.0, 3.5, 4.0}, {3.0, 3.2, 3.4}, and {3.0, 3.1, 3.2} shown with dashed lines. This shows that even a small amount of magnification diversity is beneficial when the noise variance is 0.1, and for a noise variance of 0.01 the least diverse system has an expected error 15 times lower than the single magnification system. As the noise variance decreases, performance improves for the diverse systems.

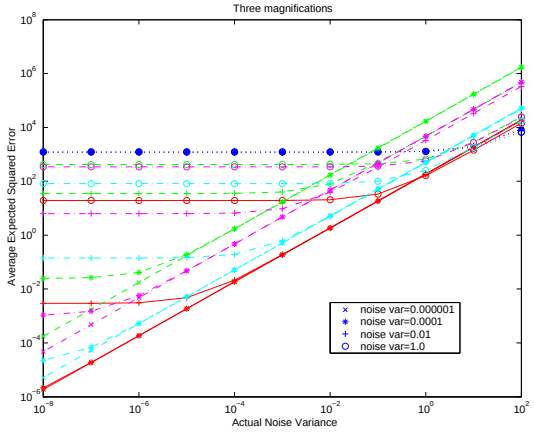

Fig. 3. Effect of small variation in q.

Actual performance on real images has a wide variability. The analysis has assumed all possible images with the specified number of bits per pixel are equally likely. However, visually interesting images usually have a higher probability of lower frequencies *i.e.*,

a $1/f^2$ type roll-off in frequency. When high contrast images are used to compare algorithms, the impact of the noise is visually less noticeable at boundary edges with large differences in pixel values. However, detail in low contrast images is far better preserved when the expected error is reduced. The images in Figure 4 are taken from an aerial view of the Bay Bridge obtained from the USC image data base. Figure 4 b shows a zoomed version of a small section of the original image in Figure 4a. The two reconstructions shown in Figures 4 c and d, use the same number of observations, but the reconstruction using magnification diversity (c) shows the low contrast structural detail more reliably than the system with single magnification factor. In both cases the noise variance used for the estimator matched the added noise variance level of 0.01.

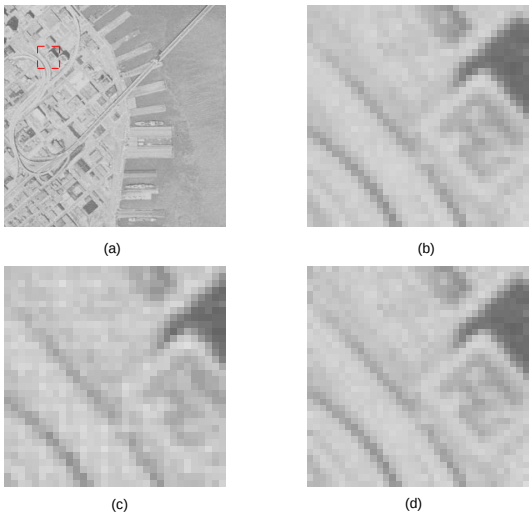

Fig. 4. Image of Bay bridge (from USC data base) (a) and 35x35 pixel tile (b). Reconstructions using a noise variance of 0.01 for (c) q=3 only and (d) q = 3, 4, 5.

5. CONCLUSIONS

This paper has extended prior work showing the benefits of magnification diversity on the improvement of super-resolution image reconstruction. Even a modest amount of diversity with less than 10% variation can lead to a reduction in average expected squared error by a factor of 15 when the noise variance is 0.01. A framework has been developed for analyzing the performance for a set of effective magnifications without assuming circulant or Toeplitz structures. The benefits of large differences in effective magnifications are partially offset by the increased impact of measurement noise as the averaging window is increased.

6. REFERENCES

[1] N. Nguyen, P. Milanfar, G.H. Golub, "A Computationally Efficient Image Superresolution Algorithm," IEEE Trans. Image Proc., vol. 10, no. 4, pp. 573-583, April 2001.

[2] Michael K. Ng and Nirmal K. Bose, "Mathematical Analysis of Super-Resolution methodology", IEEE Signal Processing Magazine, vol. 20,No.3,pp 62-74, May, 2003.

[3] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Multi-frame Super-resolution", IEEE Trans. Image Proc., vol. 13, no. 10, pp. 1327-1344 , October 2004.

[4] J. Tanida *et. al.*, "Thin Observation Module by Bound Optics (TOMBO): Concept and experimental verification," Appl. Optics-IP, vol. 40, no. 11, p. 1806-1813, Apr. 2001.

[5] Y. Kitamura *et. al.*, "Reconstruction of a high resolution image on a compound-eye image capturing system", Appl. Optics, vol 43, No. 8, p 1719-1727, 10 March 2004.

[6] M. P. Christensen, M. W. Haney, D. Rajan, S. L. Wood, S. C. Douglas, "PANOPTES: a Thin Agile Multi-Resolution Imaging Sensor," GOMACTech '05,21.5, April 2005.

[7] M. P. Christensen, V. Bhakta, D. Rajan, S. C. Douglas, S. L. Wood,"Adaptive flat multi-resolution multiplexed computational imaging architecture utilizing micro-mirror arrays to steer sub-imager field-of-views," to appear Applied Optics, 2006.

[8] S. L. Wood, B. J. Smithson, M. P. Christensen, D. Rajan, "Performance of a MVE algorithm for Compound Eye Image Reconstruction Using Lens Diversity", Proc ICASSP 2005, Philadelphia, Pa, March 2005.

[9] H-B. Lan, S. L. Wood, M. P. Christensen, D. Rajan, "Benefits of optical system diversity for multiplexed image reconstruction," to appear Applied Optics, 2006.

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

[11] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," IEEE Trans. Image Proc., vol. 6, no. 12, pp. 1646-1658, Dec. 1997.

[12] S. Baker and T. Kanade, "Limits on super-resolution and how to break them", IEEE Trans. Patt. Anal. Machine Intell.,vol. 24, no. 9, pp. 1167-1183, Sept. 2002.

[13] Z, Lin, H-Y, Shum, "Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation", IEEE Trans. PAMI, Vol. 26, no. 1, pp 83-97, January 2004.

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

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

[16] T. Kailath, Linear Systems, New Jersey: Prentice-Hall, 1980.

[17] A.K. Jain, Fundamentals of Digital Image Processing, Chapter2, Prentice Hall, 1989.

[18] 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.

[19] R. M. Gray, "On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices", IEEE Trans. Inf. Thy., vol IT-18, no.6, pp725-730, November 1972.