

# EDGE DETECTION PERFORMANCE IN SUPER-RESOLUTION IMAGE RECONSTRUCTION FROM CAMERA ARRAYS

*Sally L. Wood, Hsueh-Ban Lan\**

Santa Clara University  
Electrical Engineering Department  
Santa Clara, California, 95053  
swood@scu.edu, benh\_6\_lan@yahoo.com

*Marc P. Christensen, Dinesh Rajan*

Southern Methodist University  
Electrical Engineering Department  
Dallas, Texas, 75275  
mpc@engr.smu.edu, rajand@engr.smu.edu

## ABSTRACT

Previous work has shown that for super-resolution image reconstruction from low resolution images, image acquisition with a diverse optical system improves reconstructed image quality as measured by the expected and actual mean squared error. However, other measures of image fidelity should also be considered. An alternative performance measure might be based on edge errors, since edges are often the first step in more complex image analysis for both image processing systems and biological systems. This paper explores the behavior of edge errors and intensity errors for super-resolution image reconstruction applications in which ill-posed inversions may cause the actual mean squared error to be highly dependent of image content and thus poorly predicted by the expected mean squared error.

**Index Terms**— super resolution image reconstruction, optical system diversity, minimum variance estimator

## 1. INTRODUCTION

Super-resolution image reconstruction attempts to compute a higher resolution image using information from multiple lower resolution images (see e.g. [1]). Each individual low resolution image is captured with a slightly different projection of the 3D world. The low resolution images may come from a fixed camera array (e.g. [2]), a steerable array [3], or an image sequence from a single moveable camera (e.g. [4]). Successful algorithms could be used to create cameras based on computational imaging which might include low cost high resolution cameras or flat form-factor cameras. This image reconstruction problem is often ill-posed and improvements in performance can be achieved through diversity in the acquisition of an array of low resolution images [5, 6].

Although the average expected squared pixel error is a commonly used performance metric, it is not always the best

indicator of image quality or relevant information content. For example, adding a constant to all image values will change the brightness of an image and increase the squared error significantly. However, it will not change a viewer's perception of an image much, and many algorithms that extract information from images will not be sensitive to a constant offset. In addition, the expected mean squared error is not always a good predictor of the actual mean squared error. For ill-posed inversion problems the specific image content has a significant effect on the actual mean squared error.

A simple candidate for a versatile feature measure, in addition to intensity, is an edge element. Simple edge detection is often a first step of complex image processing tasks, and edge detection is also a first level of visual processing in biological systems. In this paper we compare the expected mean squared error to the actual mean squared error in a variety of images, and then also compare the edge detection performance using a basic Sobel edge detector (e.g. [7]). Results are compared for arrays of identical lens systems with identifiable null spaces and arrays with diverse magnification or shift variant systems which provide better observability.

## 2. MATHEMATICAL MODEL

The three dimensional source of the desired image is assumed to be far enough away from the imaging system to be accurately represented by a two dimensional surface model on a portion of a plane or sphere. This continuous source,  $f(x, y)$ , is projected onto a detector surface to create  $f_C(x, y)$ , which is bandlimited by the point spread function of the optical system and scaled relative to the image source. The objective of super-resolution image reconstruction will be to recover  $f_C(x, y)$  sampled at a desired resolution of  $(\Delta x_d, \Delta y_d)$ .

The detector surface of a digital camera contains an array of rectangular detector elements with a specified shape and array spacing. The point spread function of the optics is assumed to be small compared to the detector size, so that it is the detector array that limits the resolution of the image. Let  $a(x, y)$  represent the sensitivity of a symmetric detector cen-

---

\*The authors gratefully acknowledge the support of the Defense Advanced Research Projects Agency (DARPA) through a grant (N00014-05-1-0841) with the Office of Naval Research.

tered at the origin. The response of the single detector would be  $\int \int a(x, y) f_C(x, y) dx dy$ . The convolution of  $f_C(x, y)$  with  $a(x, y)$  produces  $g_C(x, y)$ , which represents the output of a sensor element centered at any position. This function is sampled at the detector element spacing  $(\Delta x, \Delta y)$  to produce the sampled acquired image. If  $a(x, y)$  represents a rectangular detector of width  $b$  and height  $c$  which integrates uniformly over its active surface, it will have a spatial frequency response of  $A(u, v) = (\sin(\pi ub)/(\pi u))(\sin(\pi vc)/(\pi v))$ .

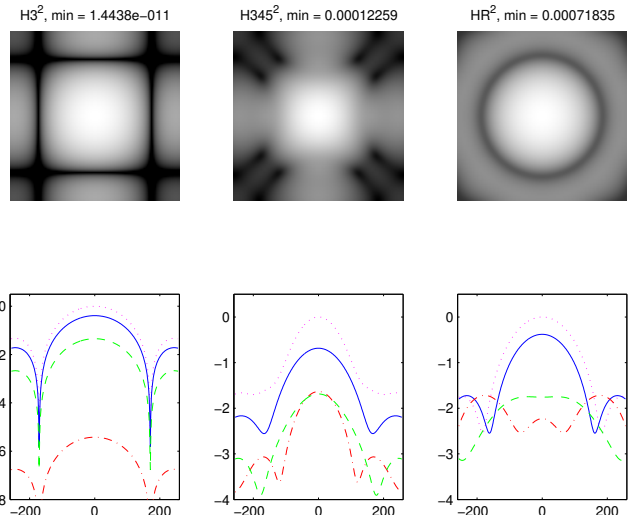
For a single array of detector elements, a sub-imager (SI), the sampling intervals  $(\Delta x, \Delta y)$  may cause aliasing of some spatial frequencies in  $f_C(x, y)$ . For increased resolution multiple images are acquired from SIs with relative translations that are fractions of the sampling intervals. Then the combined data set from all SIs will have a increased effective sampling rate, and  $f_C(x, y)$  can be estimated at frequencies higher than those obtainable from a single detector array. However, spatial frequencies for which  $A(u, v) = 0$  will not be present in any acquired image regardless of the translation of individual sensors.

For a one-dimensional image model, if images are acquired from two different arrays of SIs using detector element widths that do not have common nulled frequencies, information about all spatial frequencies will be captured in at least one set of images. For a two-dimensional model, three different widths are needed [5, 6]. Width variation for translated images may be achieved by varying the sensor size or the optical magnification. Rotation may also provide effective width variation.

Figure 1 compares the combined continuous frequency response of two systems with diversity of magnification and orientation. Arrays are translated to provide a linear magnification increase of 3. In the upper left the magnitude of the frequency response of SI arrays translated by 1/3 of the sensor width is shown as an intensity pattern of the logarithm of the response. Representative slices through the response are plotted on a logarithmic scale below. The same results are shown in the center for arrays with three different sensor sizes and on the right for arrays at three orientations. In both diverse cases the minimum combined response is much larger than the uniform array.

### 3. IMAGE RECONSTRUCTION

Let  $F$  represent the desired image  $f_C(x, y)$  sampled at the desired resolution. It will be an  $N_y$  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,  $g$ . For a traditional imaging system,  $M = N$ , but if the resolution of the observed image is lower than the desired image, then  $M < N$ .



**Fig. 1.** Continuous spatial frequency response of SI arrays: uniform, magnification diversity, and rotational diversity.

A linear mathematical model of an SI array is given by,

$$g = Hf + v, \quad (1)$$

where  $H$  is the observation matrix representing the combined effects of the optics and the detector, and the noise vector  $v$  includes uncorrelated random noise from a variety of sources. For Gaussian measurement noise and a-priori statistics with  $E[fv^T] = 0$ , the minimum variance estimate (MVE),  $\hat{f}$ , can be computed as [8],

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

Let  $\hat{R}_v = E[vv^T]$  and  $\hat{R}_v = E[vv^T]$  be the estimated noise and a-priori image variances respectively. Then the matrix  $K$  is given by

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

The covariance of the estimate error is  $\xi = E[\tilde{f}\tilde{f}^T]$  where  $\tilde{f} = f - \hat{f}$ . For an actual noise covariance  $R_v$  and an actual a-priori covariance  $P_0$ , which may be different from the estimated values, the estimation error is given by,

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

Let  $H_{0,0} = H$  represent a single SI in which the detector fields-of-view (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 can be used to increase the effective sampling rate by a factor of  $q$ . The combined  $N_{xE} \times N_{yE}$  FOV of the array will be  $q - 1$  source pixels wider and taller than the FOV of a single SI. Define a shift matrix  $Z_l$  as an  $N \times N_E$  matrix in which  $Z_l(i, j) = \delta(i - (j - l))$ . Post

multiplication of the  $H$  matrix of a SI by the correct shift will correctly position it within the combined FOV.

Let the vector  $g_{k,l}$  represent the  $(k,l)^{th}$  SI detector image. Using the measurement model of (1),

$$g_{k,l} = HZ_{kN_{x_E}+l}f + v_{k,l} = H_{k,l}f + v_{k,l}. \quad (5)$$

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. All the  $g_{k,l}$ s in (5) can be stacked in a single array in (6) where the subscript  $q$  on the combined  $H$  matrix indicates a  $q \times q$  array of SI with linear resolution reduced by a factor of  $q$ .

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

A simple reordering of  $g$  creates an image that would be obtained if  $g_C(x, y)$  were sampled at  $q$  times the actual SI sampling rate.

Using a singular value decomposition [7], the expected squared error for the MVE reconstruction can be estimated. 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. If  $R_v = \sigma^2 I_N$  and  $P_0 = p_0 I_{N_E}$ , then  $H_q P_0 H_q^T = p_0 U S S^T U^T$ . Using (4),  $\xi = V \Lambda V^T$  and the  $N_E$  diagonal elements of the diagonal matrix  $\Lambda$  are given by (7).

$$\lambda_i = \begin{cases} \frac{p_0 \sigma^2}{p_0 \sigma_i^2 + \sigma^2} & \text{for } 1 \leq i \leq N \\ p_0 & \text{for } N + 1 \leq i \leq N_E \end{cases} \quad (7)$$

For magnification diversity 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.

$$g = \begin{bmatrix} g_{q1} \\ g_{q2} \\ g_{q3} \end{bmatrix} = \begin{bmatrix} H_{q1} \\ H_{q2} \\ H_{q3} \end{bmatrix} f + v = H f + v \quad (8)$$

The performance of this diverse system can be analyzed using,

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

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

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

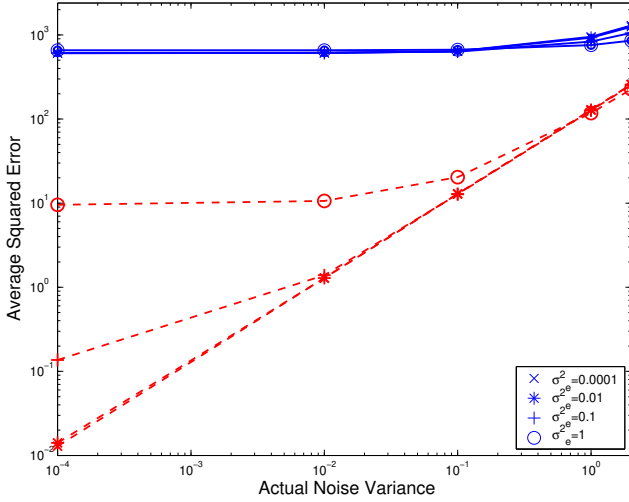
If each  $H_q$  represents a large shift invariant convolution, then the eigenvectors of all three  $H_{q_i}^T H_{q_i}$  matrices can be approximated as the same DFT vectors. Then the matrices in (10) could be added by simply adding the eigenvalues of each of the  $H_{q_i}^T H_{q_i}$  matrices scaled by the corresponding 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 as shown in Figure 1. The large desired image can be divided into smaller tiles which can be independently reconstructed using magnification diversity. For these small  $H_q$  matrices, the DFT approximation is not accurate, but exact methods can be used to invert the smaller matrices in this overdetermined set of measurements.

Figure 2 shows the expected mean squared error as a function of actual noise variance on logarithmic scales for two different imaging geometries which each use the same number of detector elements. Each was designed to improve the linear resolution by a factor of 3. One uses a uniform detector size and the other uses three different detector sizes with widths that are 1.0, 1.33 and 1.67 times the detector width of the uniform array. For each geometry results are shown for four different estimators each using a different value for the estimated noise variance.

The a-priori image covariance was set to have the pixel variance equal to the variance of a uniform distribution between 0 and 256, and it was assumed that pixel values were uncorrelated. The measurement noise can come from a number of sources and may not be well modeled by the independent Gaussian approximation. However, the number of bits per pixel should be selected so that the quantization noise does not add significant variance. Although 8-bit per pixel cameras are common, cameras with a larger number of bits can be used for high quality computational imaging. For 8-bit measurements the quantization noise variance would be 0.08 and for 16-bit measurement the variance would be  $1.3 \times 10^{-6}$ .

Actual noise variances between  $10^{-4}$  and 2 are shown. Over this range, the uniform arrays give the same result regardless of the estimated noise variance and there is no noticeable improvement as the actual measurement noise is decreased. This is due to the values of  $\lambda_i$  in (7) that are not reduced by decreasing the noise variance. In contrast, the imaging system with three different magnifications shows lower expected mse at all actual noise levels, and the expected mse is reduced as the actual noise variance is reduced until it is limited by the expected noise variance parameter used in the estimator.

The impact of these performance differences is demonstrated in reconstructions from two images. Figure 3(b) shows a segment of an aerial image of the Oakland Bay Bridge from the USC SIPI Image Database at <http://sipi.usc.edu/database>. The full image is shown to the left. Low resolution images were simulated with an added noise variance of  $\sigma_a^2 = 0.01$ , which is the same value used for the estimated noise variance. Figure 3(c) shows the reconstruction from the uniform



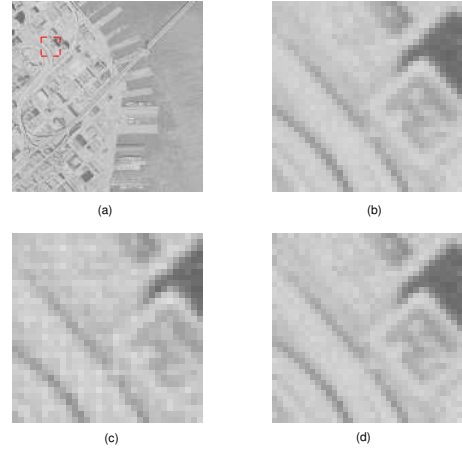
**Fig. 2.** Expected averaged squared error for estimators with  $\sigma_e^2$  and  $\sigma_a^2$ .

array in which each measurement was the average of a  $3 \times 3$  pixel area in the original image. Figure 3(d) shows results made from the same number of measurements from the array with magnification diversity as described for Figure 2. Each measurement was the average of a  $3 \times 3$  or  $4 \times 4$  or  $5 \times 5$  pixel area of the original image. The image reconstructed from the diverse array looks almost exactly like the original image and edge definition for the features is good. In contrast, the image reconstructed from the uniform array shows noticeable degradation and patterned variability. A more controlled set of soft edges is shown in Figure 4 which has edge segments at all orientations and a range of radii of curvature. The noise variance used for these simulated measurements was 0.1. The results are similar to those shown in previous image with a noticeable interference pattern in the reconstruction from the uniform array.

Figure 5(a) is an image of a vertical stripe pattern with most of its energy at a spatial frequency that is lost by the uniform sensor array. The reconstruction in Figure 5(b) has a good estimate of the average intensity value, but the stripe values and boundary positions are not correct. This error is far more significant than the lower level interference pattern. In contrast, the reconstructed image below made with the diverse array is almost indistinguishable from the original image.

#### 4. EDGE DETECTION ERROR

Many image analysis processes begin with the detection of simple edge segments as a first step to defining object boundaries or image segments. A measure of how much the variance of the image reconstruction error affects the definition of edge segments could be useful for both visual and automatic image analysis. A number of well known local edge



**Fig. 3.** (a) Bay Bridge aerial image, (b) image tile, (c) reconstruction from uniform system, (d) reconstruction from diverse magnification system

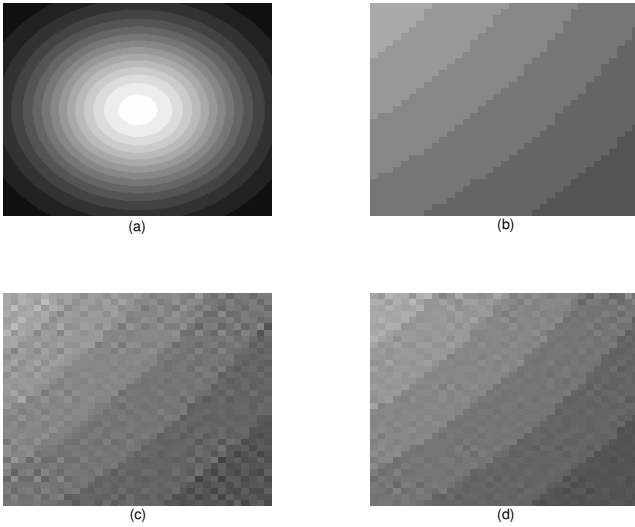
detectors use a  $3 \times 3$  pixel neighborhood to determine a local gradient magnitude and direction. The Sobel operator convolves an image with  $S_H$  and  $S_V$  as given in (11) to create two images of horizontal and vertical edge strength. Using point processing, a gradient magnitude and direction is computed for each pixel based on the horizontal and vertical edge component values for the pixel.

$$S_V = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad S_H = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (11)$$

$$\begin{aligned} \bar{s}(n)^2 &= (s(n) - \hat{s}(n))^2 \\ &= (s_x(n)^2 + s_y(n)^2) - (\hat{s}_x(n)^2 + \hat{s}_y(n)^2) \end{aligned} \quad (12)$$

The magnitude of the frequency response of the separable  $S_V$  filter is  $4|\sin(\omega_x)(1 + \cos(\omega_y))|$  where  $\omega$  is the normalized digital frequency variable. The convolution of  $S_V$  with  $\hat{f}$  will heavily attenuate the very high and very low frequencies leaving the frequencies that are considered most useful in terms of visual perception of image content. The magnitude of the frequency response of  $S_H$  will have the same form with the frequency variables interchanged.

The edge error for reconstructed images was tested using stripe patterns with slowly varying intervals and pattern changes such as Figure 5, random patterns, and a few selected images with low and high contrast. Figure 6 and Figure 7 show the average squared errors and average squared edge errors for four images. In both cases the top plots show results for an estimator based on an expected noise variance of 0.01 and the lower plots show results for a more pessimistic estimator using an expected noise variance of 1. The four images are the low contrast bay bridge image from Figure 3, a higher



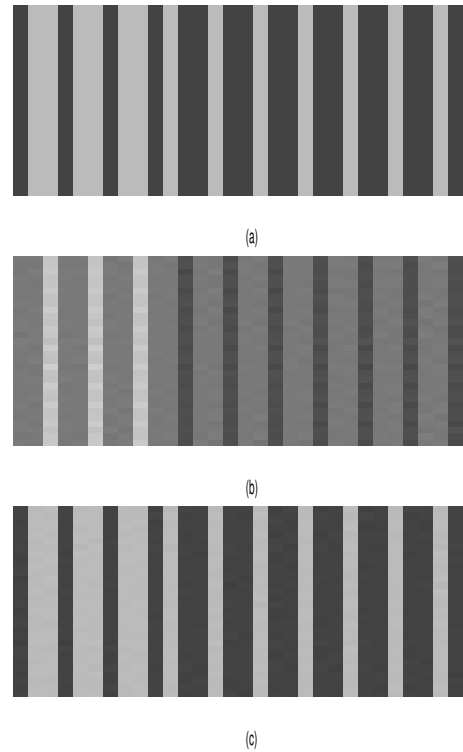
**Fig. 4.** (a) Synthetic image with ring contours, (b) image tile (c) reconstruction from uniform system, (d) reconstruction from diverse magnification system

contrast airport image from the same database, an image created with random pixel values, and the vertical stripe image from Figure 5. These results can be compared to the expected mean squared error in Figure 2.

In Figure 6(b) the reconstruction for all images by the diverse magnification system, shown in dashed lines, have approximately the same mean squared error. Since the estimated noise value is low, the mse continues to decline as the actual noise is reduced. Under these same conditions there is a large difference in the mse, shown in the solid lines, from the uniform system reconstructions. As expected from previous examples, the vertical stripe image has the largest mse, followed by the random image, the high contrast image and then the low contrast image. When the actual noise variance is reduced below  $10^{-2}$  there is little improvement in the mse. In the lower plots the more pessimistic estimator for the uniform system produces almost the same results as shown in the upper plot. However the performance of the diverse system is limited by the more cautious use of measurement data with the random image showing the worst performance and the low contrast image showing the best. For actual noise variances less than 1, the diverse magnification system significantly outperforms the uniform system. In Figure 7 the average squared edge error follows similar trends.

## 5. CONCLUSIONS

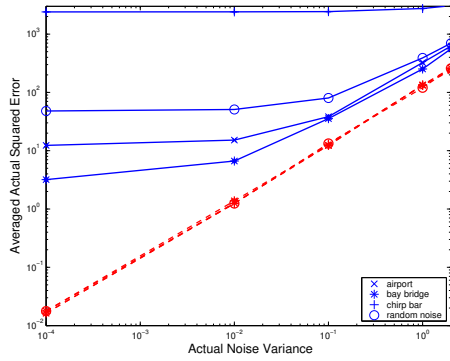
The reconstructions of a variety of images using both a uniform array and an array with magnification diversity demonstrate that as the measurement noise variance is reduced, the actual and expected noise variance in the reconstructed



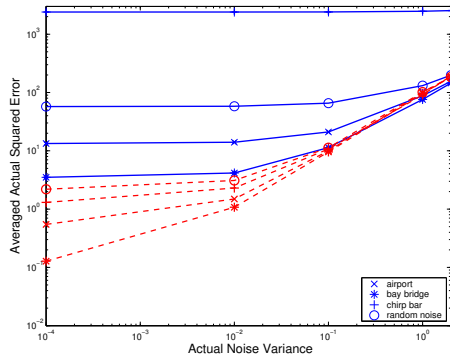
**Fig. 5.** (a) A vertical stripe synthetic image, (b) reconstruction from uniform system, (c) reconstruction from diverse magnification system

image are also reduced for the diverse array data. For the uniform array reducing the noise variance below 1 does not result in any improvement in the reconstructed image as predicted by the mathematical analysis for  $q = 3$ . In addition, depending on the specific image content, the actual mse for specific images can be far different from the expected values. This is significant because achieving noise levels lower than this floor may be essential for high quality image reconstruction allowing the confident detection of low contrast features. For the diverse system lowering the measurement noise through better detectors or longer observation times can achieve this improvement.

For both visual inspection and automated image analysis, the correct reconstruction of edges is important. Since the typical edge detection filters, such as the Sobel operator, suppress very high and very low frequencies while estimating the edges, an alternative fidelity measure using gradient magnitude was explored. It was expected that there would be image content dependent difference in this when there was significant noise variance in the reconstruction. Although some variability was observed, the behavior of the mean squared gradient error was similar to that of the mse.



(a)

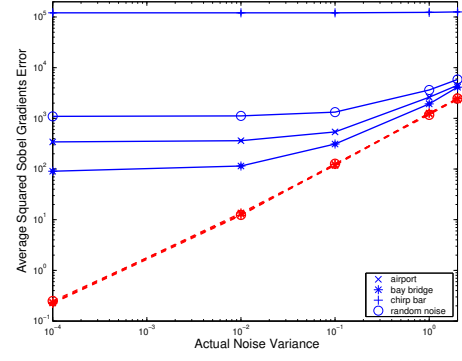


(b)

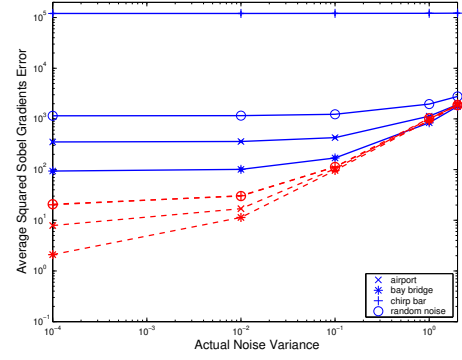
**Fig. 6.** Actual averaged squared errors for four images. In (a)  $\sigma_e^2 = 0.01$ , and in (b)  $\sigma_e^2 = 1.0$

## 6. REFERENCES

- [1] S. Chaudhuri, ed. Super-Resolution Imaging, Kluwer Academic Publishers, Boston MA, 2001.
- [2] Tanida, J. et. Al., "Thin Observation Module by Bound Optics (TOMBO): Concept and experimental verification," Applied optics-IP, Vol 40, No. 11, p1806-1813, Apr. 2001.
- [3] Christensen, M.P., V. Bhakta, D. Rajan, T. Mirani, C.C. Douglas, S.L. Wood, M.W. Haney, "Adaptive flat multiresolution multiplexed computational imaging architecture utilizing micromirror arrays to steer subimager field of views," Applied Optics, Vol 45, No. 13, May 2006.
- [4] Farsiu, S., D. Robinson, M. Elad, P. Milanfar, "Fast and robust multi-frame super-resolution," IEEE Trans. Image Proc., Vol 13, No. 10, pp1327-1344, October 2004.
- [5] Lan, H-B, S.L. Wood, M.P. Christensen, D. Rajan, "Benefits of optical system diversity for multiplexed image reconstruction," Applied Optics, Vol 45, No. 13, May 2006.



(a)



(b)

**Fig. 7.** Averaged Sobel error between desired and reconstructions for four different images. In (a)  $\sigma_e^2$  is 0.01, and in (b)  $\sigma_e^2$  is 1.0 .

- [6] Wood, S.L., H-B Lan, D. Rajan, M.P. Christensen, "Improved multiplexed image reconstruction performance through optical system diversity design," ICIP 2006.
- [7] A.K. Jain, Fundamentals of Digital Image Processing, Chapter 2, Prentice Hall, 1989.
- [8] T. Kailath, Linear Systems, New Jersey: Prentice-Hall, 1980.