STRATEGIES FOR THE USE OF STEERABLE SUB-IMAGER ARRAYS FOR IMPROVED DETECTION OF SPARSE STRUCTURES IN SUPER-RESOLUTION IMAGING

Sally L. Wood and Allen Rush
Electrical Engineering Department
Santa Clara University
Santa Clara, California 95053

ABSTRACT
In recent years research in super-resolution imaging has resulted in improved algorithms and efficient implementations for image reconstruction and image registration from multiple low resolution aliased images. However, a number of problems remain with respect to performance measures and models of noise, measurements, and image sources. As the density of pixels in typical commercial cameras has steadily increased, it is no longer safe to assume that the resolution spot size of the optics is significantly smaller than the individual pixel sensor size. In addition, the smaller sized pixel sensors typically have a higher noise level which degrades performance of reconstruction algorithms measured by PSNR or MSE. Other performance measures may be better suited to detection applications in which the source model is sparse. This paper explores the use of steerable sub-imager arrays with controlled overlapping fields-of-view to strategically reduce noise as well as increase resolution for improved detection of structures in a continuous source model.

Index Terms— super-resolution, back-projection, ART, object-detection

1. INTRODUCTION
Using computational imaging methods, a set of low resolution (LR) images acquired with sub-pixel lateral shift offsets can be used to estimate a high resolution (HR) image without the blur and aliasing due to the LR pixel array. Over the past decade a number of super-resolution camera designs and reconstruction methods have been investigated theoretically and explored in practice. Reconstruction algorithms were developed to combine information from the set of blurred and aliased LR images [1,2], and the computational burden was reduced with efficient implementation methods [3,4].

A-priori knowledge or accurate estimation of the relative shift offsets between the LR images is critical to the success of super-resolution methods, and image registration methods using the LR images have been developed [5-8]. The set of LR images may be acquired in sequence from a single camera in motion or simultaneously from multiple LR imaging devices. In the former case it is assumed that objects in the scene are motionless during the image sequence acquisition and that precise information about the camera motion is available for accurate estimation of shift values. In the latter case it is assumed that all the imagers are identical or that any differences between them are accurately known and that the relative positions of all the imagers are known.

An early example of the use of multiple imagers is the compound eye design of the Thin Observation Module by Bound Optics (TOMBO) [9-11]. Instead of a traditional camera in which a single lens focuses an image on a full sensor array, the sensor array of the TOMBO design is subdivided into a number of smaller arrays, each of which has its own small lens. Thus the field of view of the desired high resolution image is projected onto several smaller sub-arrays of pixels producing the LR image set. A fixed geometry of the lens and sub-array units determines the resolution improvement possible and the operational range which will produce the desired sub-pixel image shifts.

The Processing Arrays of Nyquist-limited Observations to Produce a Thin Electro-optic Sensor (PANOPTES) is a more flexible system developed by Christensen et al. [12]. In this design an array of micro-mirrors dynamically adjusts the fields-of-view for small sensor arrays as shown in Figure 1. A unit consisting of a small sensor array and mirrors is...
called a sub-imager (SI). One objective of this system is to achieve a HR camera with a flat form factor design which would avoid use restrictions of a traditional more boxy camera design with a longer distance between the lens and the imaging array. A second objective is to develop a versatile imaging system capable of dynamically allocating sensor resources to provide higher resolution in areas of interest at the expense of lower resolution in other areas.

Restoration of information at spatial frequencies heavily attenuated by the individual pixel sensors is a significant problem for super-resolution image reconstruction. Even if measurement noise can be significantly reduced, this limits improvement in performance in terms of the expected squared error of the reconstructed image. Previous analysis [13, 14] has shown that restoration of spatial frequencies heavily attenuated by uniform SI arrays is greatly improved by the use of SI arrays with some diversity in magnification or orientation. Then spatial frequencies heavily attenuated by one SI array may be observed more strongly in another SI array. In addition to providing better information about all spatial frequencies, diversity can also reduce the computation required. For uniform SI arrays, larger sensor arrays capture more information, but also increase the computational burden. Circulant approximations can reduce this burden [4], but this assumption may be restrictive. However, using overdetermined diverse SI arrays allows local computation over small image sub-tiles.

The measurement noise level in the LR images determines the amount or resolution improvement that can be reliably computed [13]. Noise from typical CCD cameras is not well modeled with a simple Gaussian distribution, and a large component of the noise has a Poisson distribution [15]. Averaging multiple images from each sub-imager can reduce the relative noise level and allow increased spatial resolution improvement.

Super-resolution performance is still limited by a number of factors [16]. Image registration errors are usually not well modeled although they can cause significant errors in the reconstructions. The mean squared error (MSE) or PSNR are widely used although they may not be the best performance measures. In addition, the most effective use of a-priori model information is not clear. In this paper, the effect of registration errors on MSE is demonstrated for a variety of SI geometries. For image sources modeled by a sparse distribution of isolated points, approximate reconstruction methods used in computed tomography [17] are shown to be less sensitive to noise and position errors although they do not result in the best MSE or PSNR.

2. MODELS

A typical mathematical model for super-resolution image reconstruction assumes that the desired high resolution source is a sampled image at that resolution. The desired HR image pixels, \( x \), are blurred and down sampled due to observation by a lower resolution sensor array to produce \( y \). Assume that all image arrays are stored by rows in vectors. Then one LR image, \( y_k \) would be modeled by

\[
y_k = H_k x + v
\]

where the vector \( v \) contains the additive measurement noise. Because of the resolution difference, the \( k^{th} \) LR image observation partial matrix \( H_k \) has many more columns than rows. By stacking several LR images into a single vector, \( y \), the combined measurement equation for all LR images can be written as

\[
y = H x + v
\]

and the minimum variance of error reconstruction can be written as

\[
x_{E} = K y
\]

where

\[
K = P_0 H^T (H P_0 H^T + R_v)^{-1}
\]

is one expression for the reconstruction matrix. This reconstruction matrix includes the alignment and interpolation of the LR images and the restoration required to reduce the blur due to the larger LR pixel size. For any reconstruction matrix \( K \), the covariance for the estimate error ( \( x - x_E \) ) is given by

\[
\xi = (I - KH) P_0 (I - KH)^T + KR_v K^T.
\]

If no limitation on the frequency content of the image source is assumed and the source is assumed to be continuous rather than sampled, then the objective is to reconstruct an image of that source as if it had been captured at the desired
resolution. If that desired resolution is too low, the reconstructed image will have aliasing characteristic of that resolution, but if the desired resolution is high enough, all the aliasing from the LR images will be resolved in the reconstruction process. Using that model, the reconstruction matrix is given by

\[ K = R_{xy} (R_{yy})^{-1} \]

where the correlations are computed based on the continuous image source.

3. SOURCES OF ERRORS IN RECONSTRUCTED IMAGES

Although reconstruction using (6) will minimize the MSE of the reconstructed image, this method may produce unsatisfactory results. Previous work [13, 14] showed that when multiple LR images are captured by SIs with the same magnification and orientation, spatial frequencies which are averaged to zero by the individual pixel sensors create an ill-conditioned observation partial matrix, and the expected error for the resulting reconstruction can not be arbitrarily reduced by lowering the measurement noise. However, the actual error in a specific reconstruction will depend on the source energy at spatial frequencies that are highly attenuated.

Additional increases in actual and expected reconstruction error may be caused by inaccurate values of the sub-pixel relative offsets of the LR images. Shift offset errors may be due to calibration or estimation of the shifts from the LR data. Frequency domain methods use the linear frequency dependence of phase differences to determine small shifts. But super-resolution algorithms are untended to correct aliased spatial frequencies in the LR images. However, the aliased and unaliased components at a specific frequency will have different phase offsets, and the individual contributions can not be determined. Accuracy of registration estimation methods will depend on the measurement noise level and the amount of aliasing.

The impact of errors in shift offset values can be computed in (5) using the actual observation partial matrix for H, and a reconstruction matrix K from (4) computed from the incorrect modeled value of H. To explore the impact of registration errors for varying SI sizes, the effect of offset errors was simulated for the one dimensional case. It was assumed that uniform SIs were used and that additional LR images could be acquired to reduce the impact of measurement noise on the reconstructions. The examples in Figure 2 compare two strategies for adding measurements. In all cases shown the desired resolution improvement is a factor of 2, so, in the noise free case, two SIs with a half pixel offset would be sufficient to correct aliased frequencies. The plots show the expected error variance of the HR estimate at the image center of a reconstruction for sets of 2, 4, 8 or 32 LR images as indicated by the marker shape shown in the figure legends. The error variance is plotted as a function of measurement noise variance for SIs with lengths of 3, 7, 15, 31, and 63 pixels indicated respectively by plot colors of blue, green, red, cyan, and magenta.

Figure 2: Expected error for several SI architectures as a function of measurement noise variance.
increase in spatial frequency resolution which reduced the width of the null due to sensor averaging. However since the desired spatial resolution of the HR image was only twice that of the LR image, using more than 2 LR images did not provide significant improvement. For the case of high measurement noise the number of pixels in the array had less impact, but increasing the number of LR images improved performance for all widths due to noise reduction from averaging. These results are consistent with mathematical predictions from (5) using (6) if the singular value decomposition is applied to write the observation partial matrix as \( H = USV^T \). The singular values on the main diagonal of \( S \) relative to the noise variance will determine the expected error variance of the HR image.

Figure 2b shows results similar to Figure 2a for fixed SI spacing determined by the desired spatial resolution increase. When more than two LR images were used, half had a half pixel offset from the first image and the rest were aligned with the first image. Thus, averages of multiple LR images at two specific offsets effectively provided two LR images with reduced noise. In the high and low noise regions these results are very similar to the results of Figure 2a, but in the middle range noise values, this set of offsets produced better results than those shown in Figure 2a.

The benefit of increased SI size is reduced for low noise measurements when there is uncertainty about the exact offset value for steerable or randomly positioned SI geometries. For reference the dashed lines of Figures 3a and 3b show the same results shown in Figures 2a and 2b when there is no position error, and the marker shapes and plot colors indicate the same relationship to the number and size of the SIs as was used in Figure 2. The solid lines show the expected variance when SI positions have random offset errors with a standard deviation of 0.1 pixel width. In both the uniform SI spacing case and the fixed spacing case, the largest increases in error variance for the HR estimate occurred for cases where that variance had been reduced most by reducing measurement noise or increasing number of pixels in an SI. The results in Figure 3 are typical of other values of offset uncertainty, and when the standard deviation of the offset variation is reduced to 0.05 there is still little benefit achieved by increasing the SI size above 15.

The effect of registration error is important for all super-resolution algorithms because, except for fixed and calibrated array structures, registration errors [5-7] due to aliased LR image content must be considered. These results also have implications with respect to selecting SI sizes. A smaller SI size is more flexible for variable resource allocation and local computation. But a larger number of smaller sized SIs means that more shifts must be estimated and more error could be generated.

\[
\mathbf{x}_H = H^T \mathbf{y} = H^T (H \mathbf{x} + \mathbf{v})
\]
From (4) it can be seen that this result will be similar to the least squares estimate for the very high noise case where $R_{vw} \gg H P_0 H^T$ and $P_0$ is a scaled diagonal matrix. The Algebraic Reconstruction Technique (ART) is an iterative method based on the back-projection operation. Let $x^i$ be the $i^{th}$ estimate of the full image $x$. For each individual scalar observation, $y_i$, the predicted observation error, $\Delta y_i$, is computed using $h_i^T$, the $i^{th}$ row of $H$. This observation error is back-projected to form the next estimate of $x$ as shown in (8).

$$\Delta y_i = y_i - h_i^T x^i$$

$$x^{i+1} = x^i + (\Delta y_i h_i) / \| h_i \|^2$$

After all the observations have been used, the method cycles through the same observation data multiple times. In the noise-free case this method would converge to the pseudo inverse solution, but typically the method uses a small number of cycles through the data set.

For the super-resolution application to sparse source models, this approach can provide reasonable detection with a low computational cost. Although the MSE of the reconstructed image would be higher than the MSE using (6), the result would be much less sensitive to noise in the detection of the source objects. This approach has the additional advantage that it is easy to accommodate variable spatial resolution in the reconstructed image due to dynamic resource allocation.

This method is demonstrated with a source model containing varying spatial frequencies and varying contrast, and also with a source model with a small number of randomly placed objects of small size. In both cases the linear spatial resolution of the source model definition is 4 times the spatial resolution of the desired HR result, and LR images have a reduction in linear spatial resolution by a factor of 4 compared to the desired HR result.

Figure 4a shows the image source as it would appear if captured at the desired resolution, and Figure 4b shows the image source at the desired resolution blurred by the LR pixel size. There is no aliasing at this spatial resolution, but there is attenuation at some spatial frequencies due to pixel sensor averaging. Figure 5 uses pseudo-color to show the 16 LR images with aliasing, and the interpolated LR images without restoration of the pixel blur. Figure 6 shows the same result with a small amount of added noise.

Figure 4: (a) Image source as it would appear if captured at the desired resolution. (b) at the desired resolution if captured by the LR pixel sensors.

Figure 5: Noise-free LR images showing aliasing in pseudo-color.
Figure 6: Noisy LR images showing aliasing in pseudo-color.

Figure 7a compares the reconstruction from noisy and noise-free LR images using the minimum MSE method with a 32x32 reconstruction filter with an appropriate estimate of the noise level. If a very low noise level is assumed, there is no recognizable structure in the reconstruction from the noisy data even though the noise level is very low. In Figure 7b the back-projected reconstruction shows almost no difference between the noise-free and noisy data. The minimum MSE method produces a similar result at lower amplitude when a very high estimate of the noise level is used.

Figure 8a shows a source model image for a sparse set of small objects, and Figure 8b shows 4 of the LR images and the interpolated image in the format of Figure 5. Figure 9 shows reconstruction results similar to Figure 7.

Figure 7: Image reconstruction from noise-free and noisy LR images (a) using MSE and (b) using backprojection.

Figure 8: (a) Sparse source model and (b) 4 of the 16 LR images and interpolated image at desired resolution.
5. SUMMARY

These results show that the minimum MSE super-resolution image reconstruction method can give poor results when measurement noise is present even when good estimates of the noise level are used in the estimator. When the shift offset values are not known precisely, additional errors in the MSE result are due to the use of the incorrect observation partial matrix when computing the reconstruction matrix. Use of approximate methods based on the back-projection operation reduces sensitivity to noise.

REFERENCES


