Preconditioners for Super-resolution Reconstruction*

Nhat Nguyen† Peyman Milanfar‡ Gene Golub†

Copper Mountain Iterative Methods Conference
April 3, 1998

1 Introduction

In super-resolution reconstruction we would like to reconstruct a high resolution image from a given set of low resolution frames. These low resolution frames represent different 'looks' at the same scene. Each of these frames must contribute new information which we can use to interpolate subpixel values. To get different view of the same scene, these low resolution frames must record some scene motions. These scene motions can be due to motion in the imaging system (e.g. satellite images) or motions within the scene itself (e.g. surveillance camera). If these scene motions can be estimated within subpixel accuracy, we would have information to interpolate subpixel information to recontruct a high resolution image.

Previous work on super-resolution reconstruction [1-4] has not addressed adequately the computational issues for this problem. Reconstructing a high resolution image with millions of pixels is computationally intensive. With efficient preconditioners, super-resolution reconstruction will be a more tractable process.

2 The Model

In our model, we represent each low resolution image as a noisy, down-sampled version of the high resolution image which has been shifted and blurred [2]. More formally,

$$\mathbf{b}_k = D_k C_k F_k \mathbf{x} + \mathbf{n}_k, 1 \leq k \leq p,$$

*The work in this report was supported in part by the National Science Foundation Grant CCR-9505393.

†Scientific Computing and Computational Mathematics Program, Stanford University, Stanford CA 94305-9025, {nguyen,golub}@sccm.stanford.edu.

‡Applied Electromagnetics Lab, SRI International, M/S 404-69, 333 Ravenswood Avenue, Menlo Park, CA 94025, milanfar@unix.sri.com

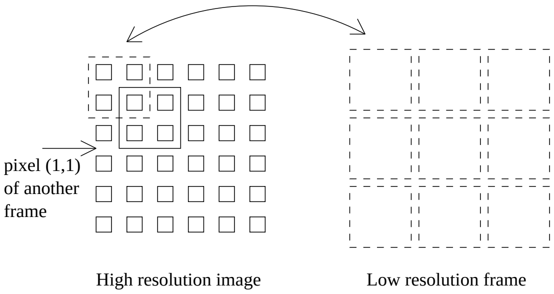

Figure 1: The Model

where p is the number of low resolution frames, $\mathbf{b}_k$ is the raster-scan vector representing the kth low resolution image, $\mathbf{x}$ is the vector representing the high resolution image, F_k is the shift matrix which represents the relative motions between frames k and a reference frame, C_k is a blur matrix, D_k is a uniform down-sampling matrix, and $\mathbf{n}_k$ is the vector representing additive white noise. We assume global translational motions and that the matrices F_k are known or can be estimated by some image registration algorithms, e.g. [5]. We will also assume that the point-spread function which generates the blurring operators C_k is known and spatially shift invariant.

Figure 1 illustrates our model conceptually. A pixel value in a low resolution frame is an "average" value over a box of pixels in the high resolution image. In the figure, the (1,1) pixel of the low resolution frame to the right is "averaged" over the dashed box while the (1,1) pixel in another frame is "averaged" over the solid box. Each low resolution frame contributes different information about the high resolution image. Super-resolution fuses information from all the given frames to reconstruct the high resolution image.

3 Preconditioners for the Regularized System

Note that the matrices C_k are derived from the discretization of a compact operator, so the given system is ill-conditioned [6]. Thus, even small measurement errors or noise in the data can result in gross error in approximations to $\mathbf{x}$ when the given system is solved directly. To obtain a reasonable estimate for $\mathbf{x}$, we regularized the problem as

$$(HH^T + \alpha LL^T)\mathbf{z} = \hat{\mathbf{b}}, H^T\mathbf{z} = \mathbf{x},$$

where $H = \begin{bmatrix} D_1C_1F_1 \\ \vdots \\ D_pC_pF_p \end{bmatrix}$, and $\hat{\mathbf{b}}$ is the given noisy data. In this formulation, L serves as a regularization matrix, and the new system is better conditioned. While a simple and effective regularization matrix can be the identity I, L can also incorporate some prior knowledge of the problem, e.g. degree of smoothness

[7]. The regularization parameter α is non-negative. A larger α corresponds to a better conditioned system, but the new system is also farther away from the original system we wish to solve. There are numerous ways to determine α, e.g. discrepancy principle, L-curve, GCV, etc. (cf [8]); we will assume α is known.

The system we wish to solve has some inherent Toeplitz structure by our assumption that the point spread function is spatially shift invariant. Using various re-orderings of the rows and columns of the system, we obtain block matrices with Toeplitz blocks. If the regularization matrix L also has Toeplitz structure, we then can construct block preconditioners with circulant blocks for the regularized system [9]. These preconditioners are easy to construct from the given matrix. They are also computationally inexpensive to apply, requiring only 2 FFTs and the solution of a banded system. Furthermore, there are theoretical results on the clustering of eigenvalues of the preconditioned system around 1. The following numerical example/experiment illustrates an example of super-resolution reconstruction and the effectiveness of circulant preconditioners for this problem.

4 Experiment

An example of super-resolution reconstruction is illustrated in Figure 2. We generated 12 low resolution frames by shifting, blurring, and down-sampling the high resolution image in the top left corner. The low resolution frames have dimensions 43×43 pixels, while the high resolution image is 172×172 pixels. We added additive Gaussian white noise with SNR = 16.7 to all low resolution frames. One of the low resolution frames is shown in the top right corner of the figure. The bottom left corner shows the result of bicubic interpolation on this low resolution frame. Using the identity as the regularization matrix with $\alpha = 0.0453$, the bottom right corner shows the result of preconditioned conjugate gradient after 2 iterations. In our experience with super-resolution reconstruction, a satisfactory approximate solution can be obtained within a few iterations of conjugate gradient.

References

[1] R. C. Hardie, T. R. Tuinstra, J. Bognar, K. J. Barnard, and E. Armstrong, *High Resolution Image Reconstruction From Digital Video Global and Non-Global Scene Motion*, in ICIP 1997.

[2] M. Elad, *Super-Resolution Reconstruction of Images*, Ph. D. Thesis - The Technion - Israel Institute of Technology, December 1996.

[3] R. S. Schultz and R. L. Stevenson, *Extraction of High-Resolution Frames from Video Sequences*, IEEE Trans. Image Processing, Vol. 5, pp. 996-1011, June 1996.

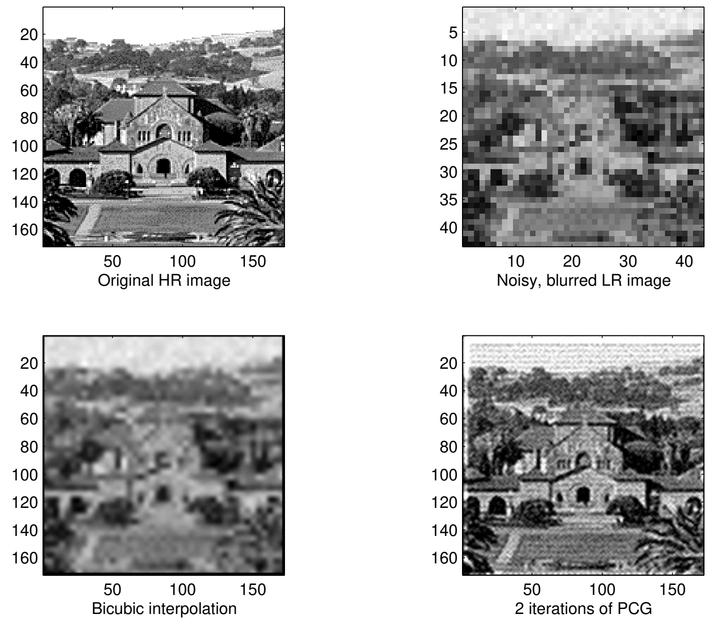

Figure 2: Stanford Quad

[4] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, *Super-resolved Surface Reconstruction from Multiple Images*, NASA Technical Report FIA-94-12, NASA Ames Research Center, Moffett Field CA, December 1994.

[5] B. K. P. Horn and B. G. Schunck, *Determining Optical Flow*, Artificial Intelligence, Vol. 17, pp. 185-204, 1981.

[6] H. Andrews and B. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[7] M. Hanke and P. C. Hansen, *Regularization Methods for Large-Scale Problems*, Surveys on Mathematics for Industry, 3 (1993), pp. 253-315.

[8] G. H. Golub and U. von Matt, *Tikhonov Regularization for Large Scale Problems*, Technical Report SCCM-97-03, SCCM Program, Stanford University, 1997.

[9] R. H. Chan and M. K. Ng, *Conjugate Gradient Methods for Toeplitz Systems*, SIAM Review, Vol. 38(3), pp. 427-482, September 1996.