PRECONDITIONERS FOR REGULARIZED IMAGE SUPERRESOLUTION

Nhat Nguyen and Gene Golub

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

Peyman Milanfar

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

ABSTRACT

Superresolution reconstruction produces a high resolution image from a set of low resolution images. Previous work on superresolution [3, 6, 10, 12] had not adequately addressed the computational issues for this problem. In this paper, we propose efficient block circulant preconditioners for solving the regularized superresolution problem by CG. Effectiveness of our preconditioners is demonstrated with superresolution results for a simulated image sequence and a FLIR image sequence.

1. INTRODUCTION

In image superresolution 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 contributes new information which we can use to interpolate subpixel values. To get different views of the same scene, these low resolution frames must record some scene motions from frame to frame. These scene motions can be due to motions 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, superresolution is possible.

2. THE MODEL

We model each low resolution image as a noisy, uniformly down-sampled version of the high resolution image which has been shifted and blurred [6]. More formally,

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

where p is the number of available frames, $\mathbf{b}_k$ is an $N \times 1$ vector representing the kth $m \times n$ ($N = mn$ pixels) low resolution image in lexicographic order. If l is the resolution factor in each direction, $\mathbf{x}$ is an $l^2N \times 1$ vector representing the $lm \times ln$ high resolution image in lexicographic order, F_k is an $l^2N \times l^2N$ shift matrix which represents the relative motions between frames k and a reference frame, C_k is a blur matrix of size $l^2N \times l^2N$, D_k is $N \times l^2N$ uniform down-sampling matrix, and $\mathbf{n}_k$ is the $N \times 1$ vector representing additive noise. We assume global translational motions and that the matrices F_k are known or can be estimated by some image registration algorithm [2]. We will also assume that the PSF which generates the blurring operator is known and spatially invariant.

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

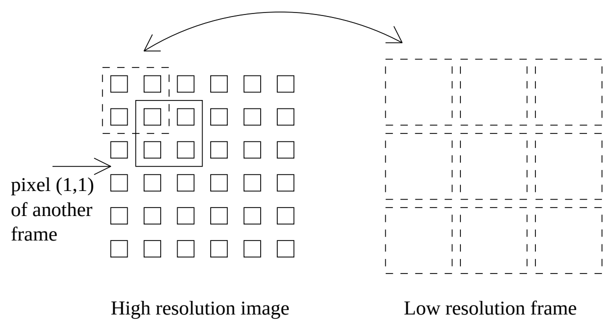

Figure 1: Superresolution model

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 an "average" over the dashed box while the (1,1) pixel in another frame an "average" over the solid box. The relative motion shift from the dashed box to the solid box is $\frac{1}{2}$ low resolution pixel down and to the right. These low resolution frames contribute different information about the high resolution image. Combining the equations in (1), we have

$$\begin{aligned} \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_p \end{bmatrix} &= \begin{bmatrix} D_1C_1F_1 \\ \vdots \\ D_pC_pF_p \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_p \end{bmatrix} \\ \mathbf{b} &= H\mathbf{x} + \mathbf{n}. \end{aligned} \tag{2}$$

If all possible combinations of subpixel horizontal and vertical shifts are available, superresolution becomes a deblurring problem. In general, this is not the case, and the system of equations above is underdetermined.

3. REGULARIZATION

The PSF is derived from the discretization of a compact operator, so H is ill-conditioned [1]. Thus, even small changes in $\mathbf{b}$ can result in wild oscillations in approximations to $\mathbf{x}$ when (2) is solved directly. To obtain a reasonable estimate for $\mathbf{x}$, we reformulate the problem as a regularized minimum norm system

$$(HH^T + \alpha R)\mathbf{z} = \mathbf{b}, \mathbf{x} = H^T\mathbf{z}. \tag{3}$$

In this formulation, R serves as a stabilization matrix, and the new system is better conditioned. While a simple and effective regularization matrix can be the identity I, R can also incorporate some prior knowledge of the problem, e.g. degree of smoothness [9]. α is the regularization parameter. 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. For the automatic calculation of regularization parameter, see [7].

4. PRECONDITIONING FOR CG

Superresolution is computationally intensive. The number of unknowns, same as the number of pixels in the high resolution image, is typically in the tens or hundreds of thousands. We must find appropriate preconditioners for (3) to accelerate CG. To exploit the structure of H, we reorder the columns of H and the elements of $\mathbf{x}$, correspondingly, as follows. We partition the high resolution image $\mathbf{x}$ into N subimages each of size $l \times l$. We enumerate the pixels in each subimage in lexicographic order 1 to l^2. The desired ordering is $q_1^{(1)}, ..., q_N^{(1)}, q_1^{(2)}, ..., q_N^{(2)}, ..., q_1^{(l^2)}, ..., q_N^{(l^2)}$, where $q_i^{(j)}$ is the jth pixel in the ith subimage. From our spatial invariance assumption of the PSF, the reordered matrix $\hat{H}$ has the following form

$$\hat{H} = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1l^2} \\ T_{21} & T_{22} & \cdots & T_{2l^2} \\ \vdots & \vdots & \ddots & \vdots \\ T_{p1} & T_{p2} & \cdots & T_{pl^2} \end{bmatrix}, \tag{4}$$

where each block T_{ij} is an $N \times N$ 'nearly'[1] Toeplitz upper band matrix. Because of the strong Toeplitz block structure in (4), we can precondition by a circulant-type preconditioner [4]. In the following subsections, we will describe efficient preconditioners for solving Toeplitz systems by CG. Extensions of these preconditioners and their convergence properties to regularized minimum norm least squares system $\hat{H}\hat{H}^T + \alpha R$ for $\hat{H}, R$ with Toeplitz *block* structure are straightforward [4].

4.1. Circulant Preconditioner

The first preconditioner, developed by Strang, completes the Toeplitz matrix T by copying the central diagonals [5]. For an upper triangular banded Toeplitz matrix T,

$$T = \begin{bmatrix} t_0 & \cdots & t_b & & \\ & \ddots & & \ddots & \\ & & t_0 & \cdots & t_b \\ & & & \ddots & \vdots \\ & & & & t_0 \end{bmatrix},$$

[1] Almost all entries along diagonals of T_{ij} are constant.

the preconditioner C_S has the form

$$C_S = \begin{bmatrix} t_0 & \cdots & \cdots & t_b & & & \\ & \ddots & & & & \ddots & \\ & & & t_0 & \cdots & \cdots & t_b \\ t_b & & & & t_0 & \cdots & t_{b-1} \\ \vdots & \ddots & & & & \ddots & \vdots \\ t_1 & \cdots & t_b & & & & t_0 \end{bmatrix}.$$

Next, we will describe our second preconditioner, developed by Hanke and Nagy [8], for banded Toeplitz matrices.

4.2. Approximate Inverse Preconditioner

Our approximate inverse preconditioner to a banded Toeplitz matrix $T_{N\times N}$ with both upper and lower bandwidths less than or equal to b is constructed as follows. First we embed T into an $(N+b)\times(N+b)$ circulant matrix C_{HN}

$$C_{HN} = \begin{bmatrix} T & T_{12} \\ T_{21} & T_{22} \end{bmatrix},$$

where

$$T_{21} = \begin{bmatrix} L & 0 & U \end{bmatrix}, T_{12} = \begin{bmatrix} U \\ 0 \\ L \end{bmatrix},$$

$$T_{22} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{b-1} \\ t_{-1} & t_0 & \cdots & t_{b-2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{-(b-1)} & t_{-(b-2)} & \cdots & t_0 \end{bmatrix}$$

$$L = \begin{bmatrix} t_b & & \\ \vdots & \ddots & \\ t_1 & \cdots & t_b \end{bmatrix}, U = \begin{bmatrix} t_{-b} & \cdots & t_{-1} \\ & \ddots & \vdots \\ & & t_{-b} \end{bmatrix}.$$

Next, we partition C_{HN}^{-1} as

$$C_{HN}^{-1} = \begin{bmatrix} M & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where M is the $N \times N$ leading principal submatrix. M is the approximate inverse preconditioner to T.

5. RESULT AND DISCUSSIONS

Note that we do not actually need to construct these preconditioners explicitly, and operations involving circulant matrices can be done efficiently by FFTs. Furthermore, we have the following result.

Theorem 1 *Let T be an upper banded Toeplitz matrix with bandwidth less than or equal to b, C_{HN} be the nonsingular extension of T, and C_S be the circulant approximation to T as above. If M is either C_S^{-1} or the $N \times N$ leading principal submatrix of C_{HN}^{-1} then*

$$MT = I + K,$$

where rank$(K) \le b$.

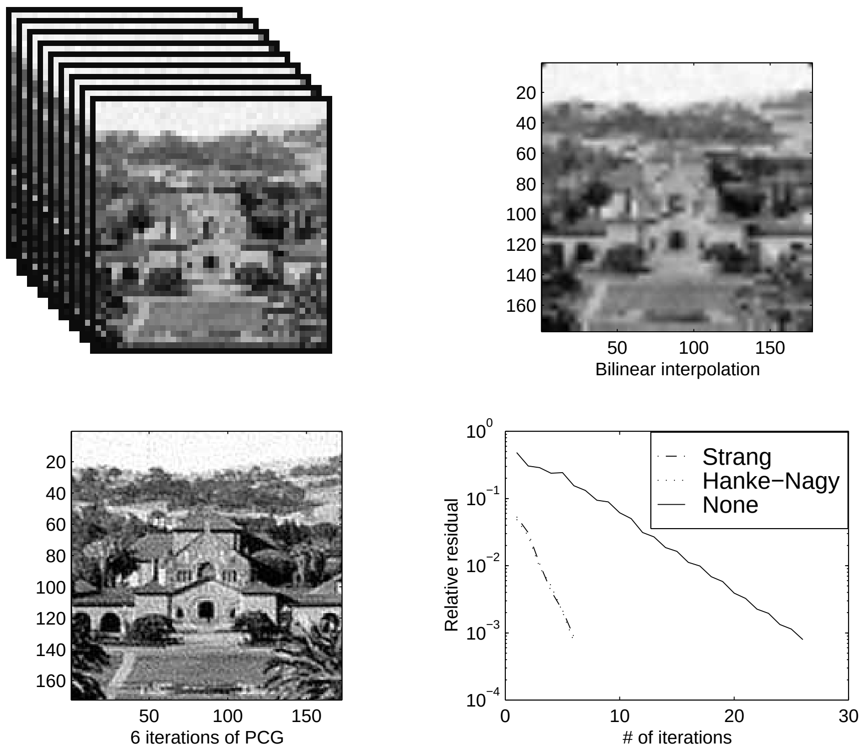

Figure 2: Superresolution on simulated sequence

The theorem above bounds the number of preconditioned CG iterations necessary for convergence. For any banded Toeplitz matrix with bandwidth b, at most b preconditioned CG iterations are needed for convergence. This result is one of the main reasons we chose our preconditioners over other circulant preconditioners, which can only claim eigenvalues of the preconditioned system "clustering" around 1 [4]. We can also bound the amount of work to solve a banded Toeplitz system by CG to $\mathcal{O}(bN\log(N))$ with Strang's circulant preconditioner and $\mathcal{O}(b(N+b)\log(N+b))$ with the approximate inverse preconditioner. In our experience, the approximate inverse preconditioner achieves slightly better performance at marginally higher cost.

6. EXPERIMENTS AND CONCLUSIONS

The first test sequence consists of artificially generated low resolution frames. We blurred a single 172×172 pixels image by a 4×4 Gaussian PSF with standard deviation of 1 and downsampled to produce 16, 43×43, low resolution frames. We added normally distributed white noise at 22.7 dB to these frames. Using 9 randomly chosen out of the complete set of 16 frames, we reconstructed the original high resolution image. Figure 2 presents the results from our superresolution algorithm. The top left portion displays the low resolution images, the top right the result of bilinearly interpolating one low resolution frames by a factor of 4 in each dimension, the bottom left the result from superresolution after 6 iterations, and the bottom right a comparison of number of CG iterations for unpreconditioned and preconditioned system to reach relative residual tolerance of 10^{-3} and with regularization parameter $\alpha = 0.01$. To reach tolerance threshold, 6 iterations of preconditioned CG are required for either preconditioners while 26 iterations are required for unpreconditioned CG.

The low resolution FLIR images in our second test sequence are provided courtesy of Brian Yasuda and the FLIR research group in the Sensors Technology Branch, Wright Laboratory, WPAFB, OH. Results using this data set are also shown in [10]. Each image is 64×64 pixels, and resolution enhancement factor of 5 is desired. The objects in the scene are stationary, and 16 frames are acquired by controlled movements of the FLIR imager. Figure 3 has similar arrangements as the previous figure. For this sequence, we again set relative residual tolerance to 10^{-3} and use regularization parameter $\alpha = 0.01$. Ten iterations are required for preconditioned CG with either preconditioners versus 28 for unpreconditioned CG to reach tolerance threshold.

For both image sequences, preconditioned CG takes about $\frac{1}{3}$ the number of iterations of unpreconditioned CG. The savings are even better for smaller regularization parameter α. Typically, we stop after 5 preconditioned CG iterations because results obtained thereafter are not significantly different visually. By these experiments, we have demonstrated that with the use of appropriate preconditioners, image superresolution is computationally much more tractable.

7. REFERENCES

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

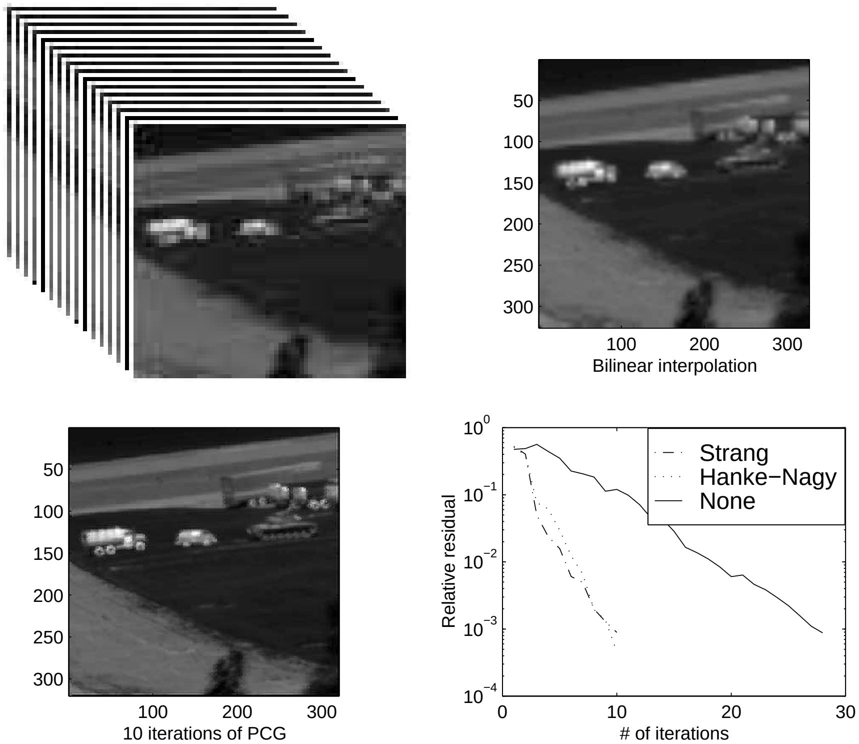

Figure 3: Superresolution on FLIR sequence

[2] L. G. Brown, *A Survey of Image Registration Techniques*, ACM Computing Survey, Vol. 24, No. 4, pp. 325-376, December 1992.

[3] 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.

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

[5] R. H. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioners*, SIAM J. Sci. Stat. Comput. Vol. 10, No. 1, pp. 104-119, January 1989.

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

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

[8] M. Hanke and J. Nagy, *Toeplitz Approximate Inverse Preconditioner for Banded Toeplitz Matrices*, Numerical Algorithms, 7(1994), pp. 183-199.

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

[10] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, *Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images*, IEEE Transactions on Image Processing, Vol. 6, No. 12, December 1997.

[11] S. Reeves and R. Mersereau, *Blur Identification by the Method of Generalized Cross-Validation*, IEEE Trans. Image Processing, Vol. 1, No. 3, pp. 301-311, July 1997.

[12] 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.