

CONJUGATE GRADIENT METHODS FOR
SUPER-RESOLUTION IMAGE
RECONSTRUCTION PROBLEMS

Guo-Dong Ye^{1 §}, Fu-Rong Lin²

^{1,2}Department of Mathematics
Shantou University

Shantou, Guangdong, 515063, P.R. CHINA

¹e-mail: guodongye@hotmail.com

²e-mail: frlin@stu.edu.cn

Abstract: This paper studies the problem of reconstructing a super-resolution image from multiple undersampled, shift, degraded frames with subpixel displacement errors. The preconditioned conjugate gradient (PCG) method with cosine transform preconditioner is efficient for solving many cases of the above problem, while it is inefficient when there are two low-resolution images taken from sensors with diagonal subpixel displacement (diagonal sensors). In this paper, we discuss the use of the conjugate gradient (CG) method with early stop criterion to the super-resolution image reconstruction problem. Numerical results show that the early stop method is very efficient for the case of diagonal sensors.

AMS Subject Classification: 26A33

Key Words: conjugate gradient, preconditioned conjugate gradient, early stop, cosine transform preconditioner

1. Introduction

Due to many limitations, hardware, for example, imaging systems often provide us with only multiple low resolution images. Multiple low resolution images of a scene are often obtained by using multiple identical image sensors which

Received: July 11, 2005

© 2005, Academic Publications Ltd.

[§]Correspondence author

are shifted relative to each other by subpixel displacements [1, 5, 6]. Super-resolution image reconstruction refers to obtaining an image at a resolution higher than that of the sensors used in recording the image [7, 10]. High-resolution image reconstruction is a special case, in which case we have enough low-resolution images with suitable subpixel displacements [9].

Super-resolution image reconstruction has many electronic image applications such as aerial or facilities surveillance, commercial and scientific imaging and so on. The high-resolution image reconstruction problem is closely related to the design of high-definition television (HDTV) and very high-definition (VHD) image sensors. CCD image sensor arrays, where each sensor consists of a rectangular subarray of sensing elements, produce discrete images whose sampling rate and resolution are determined by the physical size of the sensing elements.

Ng et al [9] used cosine transform preconditioners to precondition the linear system resulting from high-resolution image reconstruction problems. When the number of shifted low-resolution is equal to four in 2-by-2 sensor setting and these four shifted low-resolution images are shifted relative to each other by the half pixel value, the preconditioned conjugate gradient (PCG) method applied to solving the preconditioned system converges superlinearly. For the PCG methods, we refer readers to the book by Golub and Van Loan [4]. We note that under the noiseless condition, the four shifted low-resolution image are sufficient to reconstruct the high-resolution image perfectly. In [10], Ng and Sze further modified cosine transform preconditioners to handle the cases where the number of shifted low-resolution image is equal to two. Numerical results showed that those preconditioners are quite efficient for the cases where the two sensors are horizontal or vertical, while it is inefficient if the two sensors are diagonal. The preconditioners based on incomplete Cholesky factorization proposed by Lin et al [7] are also inefficient for the case of diagonal sensors.

In Section 2, we introduce the mathematical formulation of the problem and the use of conjugate gradient type methods for solving the resulted mathematical problem. In particular, we consider using the CG method with early stop criterion to the super-resolution reconstruction problem. One advantage of the early stop method is that the selection of optimal parameter is not required. Numerical results in Section 3 show that the early stop method is very efficient for the case of diagonal sensors.

2. Super-Resolution Image Reconstruction

2.1. Mathematical Model

In this subsection, we will introduce the mathematical model for the super-resolution image reconstruction. Suppose that we have m sensors, each sensor has $N_1 \times N_2$ sensing elements (pixels) and the size of each sensing element is $T_1 \times T_2$, we then have m images of resolution $N_1 \times N_2$ (low-resolution images). Our aim is to reconstruct an image of resolution $M_1 \times M_2$ (high-resolution images), where $M_1 = L_1 \times N_1$ and $M_2 = L_2 \times N_2$.

In order to have some information to resolve the high-resolution image, there are subpixel displacement between the sensors. More precisely, there exist integers $\mu_i \in [0, L_1 - 1]$, $\nu_i \in [0, L_2 - 1]$, and real numbers $\epsilon_i^x, \epsilon_i^y \in (-\frac{1}{2}, \frac{1}{2})$, such that the horizontal and vertical displacements of the i -th sensor are given by

$$d_i^x = \frac{T_1}{L_1}(\mu_i + \epsilon_i^x) \quad \text{and} \quad d_i^y = \frac{T_2}{L_2}(\nu_i + \epsilon_i^y).$$

Here (μ_i, ν_i) is the sensor position of the i -th sensor, and ϵ_i^x and ϵ_i^y denote respectively the normalized horizontal and vertical displacement errors. We note that the parameters ϵ_i^x and ϵ_i^y can be obtained by manufacturers during camera calibration. The estimation of these displacement errors was discussed in [8].

Let the positions of all sensors be denoted by

$$S = [(\mu_1, \nu_1), \dots, (\mu_m, \nu_m)].$$

We note that $S = [(\mu, \nu) : 0 \leq \mu < L_1, 0 \leq \nu < L_2]$ corresponds to the case of the *high-resolution* image reconstruction and $S \subset [(\mu, \nu) : 0 \leq \mu < L_1, 0 \leq \nu < L_2]$ corresponds to the case of *super-resolution* image reconstruction.

Let f be the original scene. The observed low-resolution image g_i is modeled by

$$g_i[n_1, n_2] = \frac{1}{T_1 T_2} \int_{T_2(n_2 - \frac{1}{2}) + d_i^y}^{T_2(n_2 + \frac{1}{2}) + d_i^y} \int_{T_1(n_1 - \frac{1}{2}) + d_i^x}^{T_1(n_1 + \frac{1}{2}) + d_i^x} f(x, y) dx dy + \eta_i[n_1, n_2] \quad (1)$$

for $n_1 = 1, 2, \dots, N_1$ and $n_2 = 1, 2, \dots, N_2$. Here η_i is the noise corresponding to the i th sensor. Similarly, the high-resolution image \mathbf{z} is modeled by

$$z[n_1, n_2] = \frac{L_1 L_2}{T_1 T_2} \int_{(n_2 - \frac{1}{2})T_2/L_2}^{(n_2 + \frac{1}{2})T_2/L_2} \int_{(n_1 - \frac{1}{2})T_1/L_1}^{(n_1 + \frac{1}{2})T_1/L_1} f(x, y) dx dy. \quad (2)$$

Let \mathbf{g}_i , $\vec{\eta}_i$, and \mathbf{z} be the corresponding vectors obtained by using a column by column ordering for g_i , η_i and z , respectively, we have

$$\mathbf{g}_i = \mathbf{H}_i \mathbf{z} + \vec{\eta}_i,$$

where \mathbf{H}_i is the blurring matrix corresponding to the i -th sensor [1].

We can formulate the reconstruction problem as finding \mathbf{z} such that

$$\sum_{i=1}^m \|\mathbf{g}_i - \mathbf{H}_i \mathbf{z}\|_2^2 \quad (3)$$

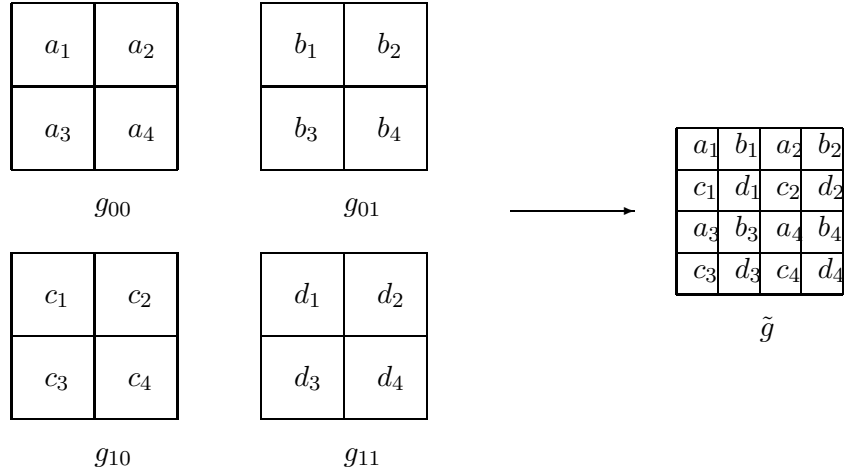
is minimized. In the following, we specify the matrices \mathbf{H}_i . Because of the blurring process, the boundary values of \mathbf{g}_i are also affected by the values of f outside the scene. We use the Neumann boundary condition, which assumes that the scene immediately outside is a reflection of the original scene at the boundary [9].

Let $d_{\mu,\nu}$ be the $\nu \times 1$ vector with zero entries except its $(\mu+1)$ -th entry being equal to 1 (for instance, $d_{1,4} = (0, 1, 0, 0)^t$). Under the Neumann boundary condition, the blurring matrices are given by

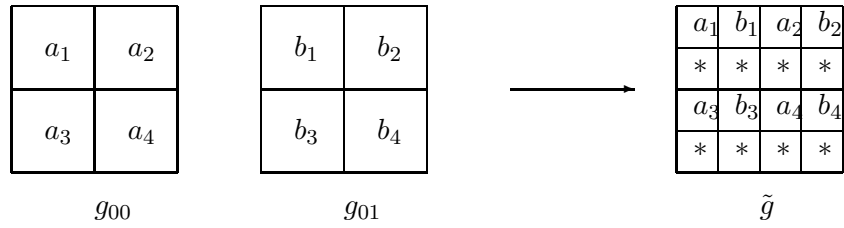
$$\mathbf{H}_i = [(\mathbf{I}_{M_2/L_2} \otimes d_{\nu_i, L_2}^t) \mathbf{H}^y(\epsilon_i^y)] \otimes [(\mathbf{I}_{M_1/L_1} \otimes d_{\mu_i, L_1}^t) \mathbf{H}^x(\epsilon_i^x)], \quad (4)$$

where $\mathbf{H}^x(\epsilon_i^x)$ and $\mathbf{H}^y(\epsilon_i^y)$ are $M_1 \times M_1$ and $M_2 \times M_2$ Toeplitz-plus-Hankel matrices, respectively:

$$\mathbf{H}^x(\epsilon_i^x) = \frac{1}{L_1} \begin{pmatrix} 1 & \cdots & 1 & h_i^{x+} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ 1 & \ddots & \ddots & \ddots & \ddots & h_i^{x+} \\ h_i^{x-} & \ddots & \ddots & \ddots & \ddots & 1 \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_i^{x-} & 1 & \cdots & 1 \end{pmatrix} + \frac{1}{L_1} \begin{pmatrix} 1 & \cdots & 1 & h_i^{x-} & & 0 \\ \vdots & \cdots & \cdots & & & \\ 1 & \cdots & & & & h_i^{x+} \\ h_i^{x-} & & & & \cdots & 1 \\ & & & \cdots & \cdots & \vdots \\ 0 & & h_i^{x+} & 1 & \cdots & 1 \end{pmatrix},$$



(a) High-resolution image reconstruction



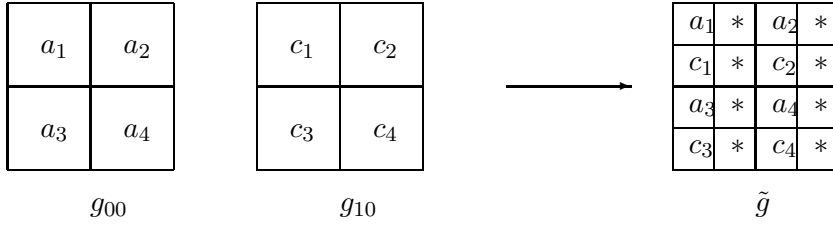
(b) High-resolution image reconstruction (horizontal sensors)

Figure 1: Observed images

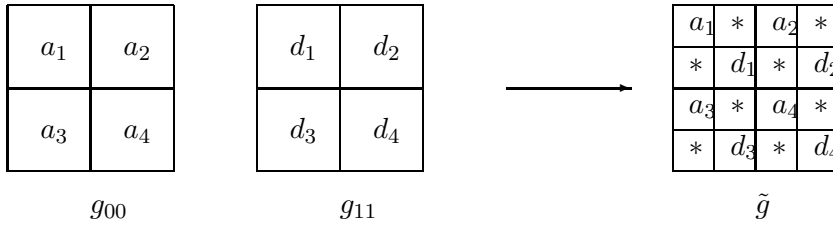
where \mathbf{L} is the discretization of regularization operator such as the identity operator or a low order differential operator and the regularization parameter α is a small positive number controlling the degree of regularity of the solution.

Using the first order differential operator and Neumann boundary conditioner, we get

$$\mathbf{L}^t \mathbf{L} = \mathbf{I}_{M_2} \otimes \mathbf{L}_{M_1} + \mathbf{L}_{M_2} \otimes \mathbf{I}_{M_1},$$



(c) High-resolution image reconstruction (vertical sensors)



(d) High-resolution image reconstruction (diagonal sensors)

Figure 1: Continuation: Observed images

where \mathbf{L}_M is the $M \times M$ matrix

$$\mathbf{L}_M = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}.$$

We let $\mathbf{H}_i^x = (\mathbf{I}_{M_1/L_1} \otimes d_{\mu_i, L_1}^t) \mathbf{H}^x(\epsilon_i^x)$ and $\mathbf{H}_i^y = (\mathbf{I}_{M_2/L_2} \otimes d_{\nu_i, L_2}^t) \mathbf{H}^y(\epsilon_i^y)$, such that \mathbf{H}_i defined in (4) has the form $\mathbf{H}_i = \mathbf{H}_i^y \otimes \mathbf{H}_i^x$. Here let $\epsilon = [(\epsilon_1^x, \epsilon_1^y), \dots, (\epsilon_m^x, \epsilon_m^y)]$, $\mathbf{g} = \sum_{i=1}^m (\mathbf{H}_i^y \otimes \mathbf{H}_i^x)^t \mathbf{g}_i$, and

$$\mathbf{A}(S, \epsilon, \alpha) = \sum_{i=1}^m [(\mathbf{H}_i^y)^t \mathbf{H}_i^y] \otimes [(\mathbf{H}_i^x)^t \mathbf{H}_i^x] + \alpha \mathbf{L}^t \mathbf{L}. \tag{6}$$

It can be seen that the optimization problem (5) is equivalent to

$$\mathbf{A}(S, \epsilon, \alpha) \mathbf{z} = \mathbf{g}. \tag{7}$$

Since (7) is often a very large system, say, if $M_1 = M_2 = 128$, then the number of unknowns is $128^2 = 16384$, only iterative methods such as conjugate

type methods can be applied. The PCG methods with cosine transform preconditioners and incomplete Cholesky factorization preconditioners have been considered by Ng et al [9] and Lin et al [7], respectively. In these PCG methods, it is crucial to choose an optimal parameter α , otherwise, the reconstructed high-resolution image may be of low quality. Numerical results show that these methods are not efficient for the case $S = [(0, 0), (1, 1)]$.

We consider apply using the CG method with early stop criterion [2] to solve (7). In this method, we can use a very small regularization parameter (say, $\alpha = 10^{-12}$) independent of the noise level of low-resolution images. We use the following stopping criterion

$$\frac{\|\mathbf{r}^{(k+1)} - \mathbf{r}^{(k)}\|_2}{\|\mathbf{r}^{(k+1)}\|_2} \leq s(\text{SNR}), \quad (8)$$

where SNR denote the signal to noise rate of the low-resolution images and $\mathbf{r}^{(k)}$ is the residual vector of the k -th iteration. In our test, we defined $s(\cdot)$ by (10) in the next section. We note that in the PCG method, an often used stopping criterion is

$$\frac{\|\mathbf{r}^{(j)}\|_2}{\|\mathbf{r}^{(0)}\|_2} \leq \epsilon, \quad (9)$$

where ϵ is a small positive number such as 10^{-6} . We like to note that the cost per iteration of the CG method is $O(M_1 M_2)$, which is considerably less than that of the PCG method with cosine transform preconditioner seen as $O(M_1 M_2 \log(M_1 M_2))$.

3. Numerical Experiments

In this section, we compare the performance of the PCG method with cosine transform preconditioner (with stopping criterion given by (9), $\epsilon = 10^{-6}$), the CG method with stopping criterion (9) ($\epsilon = 10^{-6}$), and the CG method with stopping criterion (8) with

$$s(x) = \begin{cases} 1.3700 \cdot e^{-0.07 \cdot x}, & x \geq 40, \\ -0.0217 \cdot x + 0.9510, & x \in [30, 40]. \end{cases} \quad (10)$$

In the following, we denote the three methods by PCG, CG, and CGE, respectively. Four sensor array settings of the 2×2 case are tested:

- (i) $S1 = [(0, 0), (1, 1)]$; (ii) $S2 = [(0, 0), (0, 1), (1, 0), (1, 1)]$;
- (iii) $S3 = [(0, 0), (0, 1)]$; (iv) $S4 = [(0, 0), (1, 0)]$.

S	PCG(CG)	CGE
	$\alpha = 5.0 \times 10^{-5}$	$\alpha = 1.0 \times 10^{-12}$
S1	0.0522	0.0537
S2	0.0312	0.0372
S3	0.0701	0.0909
S4	0.0617	0.0725
S	PCG(CG)	CGE
	$\alpha = 5.0 \times 10^{-3}$	$\alpha = 1.0 \times 10^{-12}$
S1	0.0638	0.0646
S2	0.0551	0.0566
S3	0.0765	0.0951
S4	0.0706	0.0803

Table 1: Relative errors of the reconstructed images where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Bridge)

Two images Bridge and Cameraman are tested. The original images, one of the low-resolution images of 50dB are shown in Figures 2(a), 2(b), 3(a) and 3(b) for $S1$. The displacement errors ϵ_i^x and ϵ_i^y are chosen randomly between -0.1 and 0.1 . In all methods, we choose zero vector as initial guesses. In the PCG and CG methods, we choose optimal regularization parameters [10]. Note that the parameters α are different for different images, cf. Table 1 and Table 4. In the CGE method, we simply choose $\alpha = 10^{-12}$ without considering the noise level of low-resolution images. We also show the observed image, reconstructed images obtained by the CGE and PCG methods. The computations are done by *Matlab* on a PC of 1100MHz CPU and 256M REM.

In Table 1 – Table 6, we show the relative errors of the reconstructed images obtained by PCG, CG, and CGE methods and the number of iteration required and also the time needed for convergence of the above three methods. Each relative error is an average of 30 tests, so are the number of iterations and the running time.

Table 2 and Table 5 show the time required for the CG, PCG and CGE to restore the high-resolution and super-resolution from low-resolutions. We observe that the running time of CGE method is greatly less than those of the PCG and CG methods except in case $S2$. From Table 3 and Table 6, we observe that except the case $S2$, the number of iterations for the CGE is also considerably less than those of the PCG and CG methods.

S	CG	PCG	CGE	S	CG	PCG	CGE
S1	8.3014	16.1708	1.6011	S1	1.6984	2.9890	0.5744
S2	13.8228	3.2042	2.3242	S2	2.4061	1.9046	0.6273
S3	12.6997	34.3585	1.4249	S3	2.4771	5.2869	0.4457
S4	15.5356	38.6972	1.5325	S4	2.5296	5.8506	0.4293

Table 2: The time (seconds) needed for CG, PCG and CGE to restore where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Bridge)

S	CG	PCG	CGE	S	CG	PCG	CGE
S1	138.1	107.3	28.1	S1	27.0	17.2	7.0
S2	160.7	11.3	26.7	S2	31.3	6.5	7.0
S3	232.4	240.5	25.3	S3	40.1	33.8	6.2
S4	231.4	240.6	27.2	S4	40.9	36.9	6.4

Table 3: Number of iterations required for convergence where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Bridge)

S	PCG(CG)	CGE
	$\alpha = 2.0 \times 10^{-4}$	$\alpha = 1.0 \times 10^{-12}$
S1	0.0397	0.0411
S2	0.0212	0.0219
S3	0.0493	0.0648
S4	0.0753	0.0914
S	PCG(CG)	CGE
	$\alpha = 5.0 \times 10^{-3}$	$\alpha = 1.0 \times 10^{-12}$
S1	0.0539	0.0558
S2	0.0460	0.0469
S3	0.0609	0.0739
S4	0.0806	0.0960

Table 4: Relative errors of the reconstructed images where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Cameraman)

S	CG	PCG	CGE	S	CG	PCG	CGE
S1	5.1295	8.9182	2.3598	S1	1.6240	2.9192	0.4545
S2	7.6409	2.3159	2.9429	S2	2.5794	1.6548	0.5782
S3	8.7515	21.2138	1.8650	S3	2.4664	5.5757	0.4375
S4	9.0768	21.4475	1.8755	S4	2.5030	5.5008	0.4054

Table 5: The time (seconds) needed for CG, PCG and CGE to restore where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Cameraman)

S	CG	PCG	CGE	S	CG	PCG	CGE
S1	86.4	61.4	39.5	S1	27.0	17.0	7.2
S2	98.4	9.1	32.8	S2	32.0	6.4	7.0
S3	140.8	134.4	27.6	S3	40.2	34.9	7.3
S4	139.8	129.5	27.0	S4	41.0	36.0	6.4

Table 6: Number of iterations required for convergence where the low-resolution images are of 50 dB (left) and 30 dB (right), respectively (Cameraman)

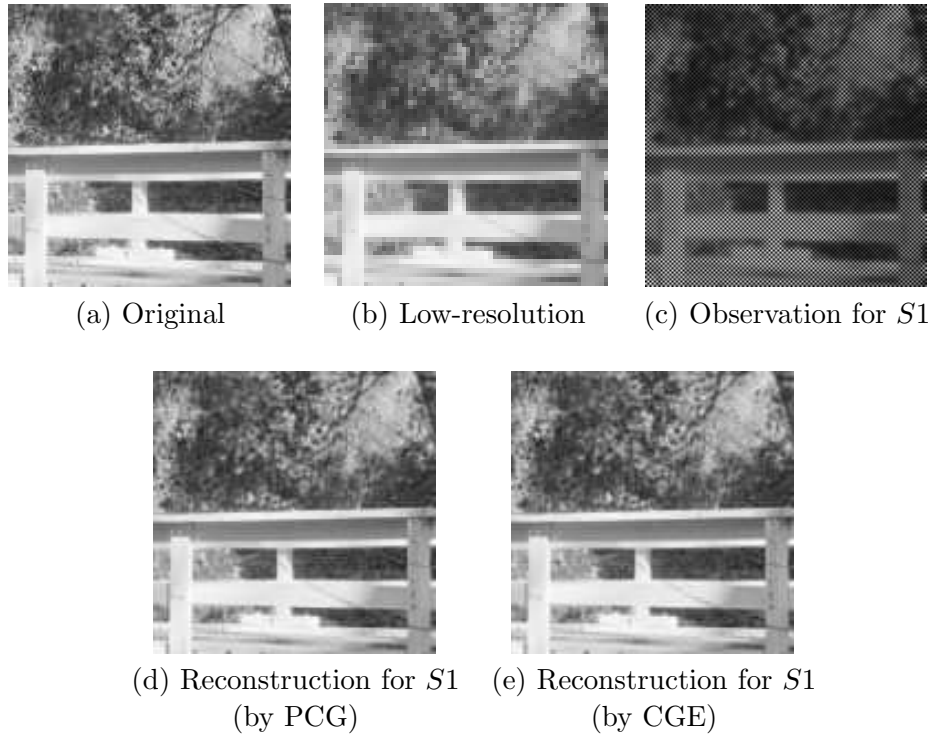


Figure 2: Images for Bridge

As for the quality of the reconstructed images, the result of the CGE method is not as good as that of the PCG (CG) method, cf. Table 1 and Table 4. However, we see that for S_1 , the relative errors of the reconstructed images of the CGE method are only slightly larger than those of the PCG method.

4. Concluding Remarks

We compare the performance of three conjugate gradient type methods for super-resolution image reconstruction problems. It can be seen that cosine transform preconditioner is well suitable for the case of high-resolution image reconstruction problems [9] and the early stop method proposed in this paper is suitable for diagonal sensors. We like to note that the method proposed in [10] is suitable for horizontal and vertical sensor arrays.

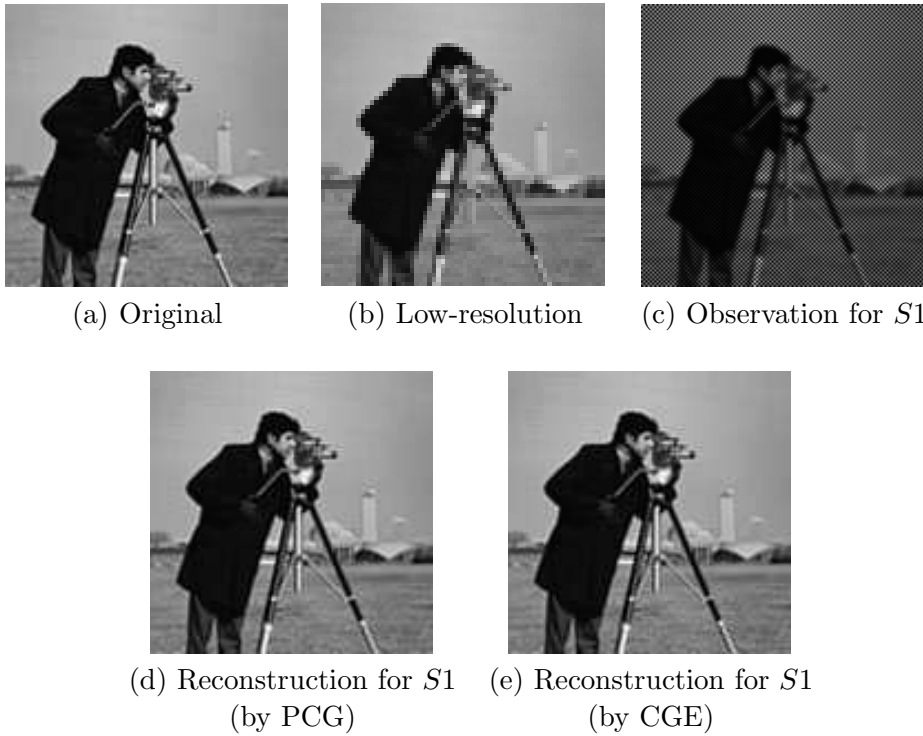


Figure 3: Images for Cameraman

Acknowledgements

Research is supported in part by the Guangdong Provincial Natural Science Foundation of P.R. China, No. 021244.

References

- [1] N.K. Bose, K.J. Boo, High-resolution image reconstruction with multisensors, *International Journal of Imaging Systems and Technology*, **9** (1998), 294-304.
- [2] H. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht (1996).
- [3] A. Forsgren, P.E. Gill, W. Murry, Computing modified Newton directions

- using a partial Cholesky factorization, *SIAM J. Sci. Comput.*, **16** (1995), 139-150.
- [4] G.H. Golub, C.F. Van Loan, *Matrix Computations*, Third Edition, The Johns Hopkins University Press (1996).
- [5] G. Kainlath, V. Odet, R. Goutte, Image resolution enhancement using subpixel camera displacement, *Signal Processing*, **26** (1992), 139-146.
- [6] T. Komatsu, K. Aizawa, T. Igarashi, T. Saito, Signal processing based method for acquiring very high resolution images with multiple cameras and its theoretical analysis, *IEE Proceedings*, **140-3**, Part I (1993), 19-25.
- [7] F.R. Lin, W.K. Ching, M.K. Ng, Preconditioning regularized least squares problems arising from high-resolution image reconstruction from low-resolution frames, *Linear Algebra and its Applications*, **391** (2004), 149-168.
- [8] M. Ng, N.K. Bose, J. Koo, Constrained total least squares computations for high resolution image reconstruction with multisensors, *International Journal of Imaging Systems and Technology*, **12** (2002), 35-42.
- [9] M. Ng, R. Chan, T. Chan, A. Yip, Cosine transform preconditioners for high resolution image reconstruction, *Linear Algebra and its Applications*, **316** (200), 89-104.
- [10] M. Ng, K.N. Sze, Preconditioned iterative methods for super-resolution image reconstruction with multisensors, In: *Symposium on Advanced Signal Processing: Algorithms, Architectures and Implementations* (Ed. Franklin Luk); *Proceedings to the SPIE*, **4116** (2000), 396-405, San Diego CA, July (2000).