BAYESIAN HIGH RESOLUTION IMAGE RECONSTRUCTION WITH INCOMPLETE MULTISENSOR LOW RESOLUTION SYSTEMS

Javier Mateos[a], Rafael Molina[a], and Aggelos K. Katsaggelos[b]*

a) Dpto. de Ciencias de la Computación e I.A., Universidad de Granada, 18071 Granada, Spain
b) Dept. of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208-3118

ABSTRACT

In this paper we consider the problem of reconstructing a high-resolution image from an incomplete set of undersampled, shifted, degraded frames with subpixel displacement errors. We derive mathematical expressions for the calculation of the maximum *a posteriori* (MAP) estimate of the high resolution image given the low resolution observed images. We also examine the role played by the prior model when an incomplete set of low resolution images is used. Finally, the proposed method is tested on real and synthetic images.

1. INTRODUCTION

Over the last two decades research has been devoted to the problem of reconstructing a high-resolution image from multiple undersampled, shifted, degraded frames with subpixel displacement errors (see [1] for a review). Most of the reported work addresses the problem of estimating an $LM \times LN$ high resolution image from at least $L \times L$ low resolution images of size $M \times N$. Kim et al. [2] explore the conditions the shifts of the $L \times L$ low resolution images have to satisfy in order to solve the high resolution problem, at least from the least square perspective. Elad and Feuer [3] study the same problem when combining Bayesian, Projection onto Convex Sets and Maximum Likelihood methodologies on high resolution problems. Baker and Kanade [4] also examine the impact of increasing the number of low resolution images, when proposing an alternative approach to the super resolution problem. However, not much work has been reported on the role played by the prior model when the system is incomplete, that is, when we have less than $L \times L$ low resolution images or when the shifts do not satisfy the conditions in [2] or [3].

In this paper we use the general framework for frequency domain multi-channel signal processing developed by Katsaggelos et al. in [5] (a formulation that was also obtained later by Bose and Boo [6] for the high resolution problem) to tackle the high resolution problem from incomplete observations. With the use of block-semi circulant matrices we show that all the matrix calculations involved in the MAP estimation can be performed in the Fourier domain. We also examine how the prior model compensates for the lack of information in the incomplete observation set.

The rest of the paper is organized as follows. The problem formulation is described in section 2. In section 3 the degradation and image models used in the Bayesian paradigm are described. The application of the Bayesian paradigm to calculate the MAP high resolution image and estimate the hyperparameters is described in section 4. Experimental results are described in section 5. Finally, section 6 concludes the paper.

*This work has been partially supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC2000-1275.

2. PROBLEM FORMULATION

Consider a camera sensor with $N_1 \times N_2$ pixels and assume we have a set of $1 \leq q \leq L \times L$ shifted images. Our aim is to reconstruct an $M_1 \times M_2$ high resolution image with $M_1 = L \times N_1$ and $M_2 = L \times N_2$, from the set of low-resolution observed images.

The low resolution sensors are shifted with respect to each other by a value proportional to $T_1/L \times T_2/L$, where $T_1 \times T_2$ is the size of each sensing element (note that if the sensors are shifted by values proportional to $T_1 \times T_2$ or $q < L \times L$ the high-resolution image reconstruction problem becomes singular). The normalized horizontal and vertical displacements may be assumed to be known (see [6, 7] for details) or unknown (see [8] for an approach where the displacements are assumed unknown and are estimated simultaneously with the high-resolution image).

Let $\mathbf{g}_{l1,l2}$ be the $(N_1 \times N_2) \times 1$ observed low resolution image acquired by the $(l1, l2)$-th sensor. Our goal is to reconstruct $\mathbf{f}$, the $(M_1 \times M_2) \times 1$ high resolution image, from a set of q low resolution images $\mathbf{g}_{l1,l2}$, with $1 \leq q \leq L^2$. We will denote by $\mathcal{I}$, the set of indices of the available low resolution images.

The process to obtain the observed low resolution image by the $(l1, l2)$-th sensor, $\mathbf{g}_{l1,l2}$, from $\mathbf{f}$ can be modeled as follows. First, $\mathbf{f}^{l1,l2}$ is obtained which represents a blurred version of the original high-resolution one, according to

$$\mathbf{f}^{l1,l2} = \mathbf{H}_{l1,l2}\mathbf{f}, \tag{1}$$

where $\mathbf{H}_{l1,l2}$ is an $(M_1 \times M_2) \times (M_1 \times M_2)$ integrating matrix that may have different forms. In this paper we use an $\mathbf{H}_{l1,l2}$ representing a linear space-invariant blurring system with impulse response

$$h_{l1,l2}(u,v) = \begin{cases} \frac{1}{L^2} & u,v = -(L-1), \ldots, 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

Let now $\mathbf{D}_{l1}$ and $\mathbf{D}_{l2}$ be the 1-D downsampling matrices defined by

$$\mathbf{D}_{l1} = \mathbf{I}_{N_1} \otimes \mathbf{e}_l^t, \qquad \mathbf{D}_{l2} = \mathbf{I}_{N_2} \otimes \mathbf{e}_l^t, \tag{3}$$

where $\mathbf{I}_{N_i}$ is the $N_i \times N_i$ identity matrix, $\mathbf{e}_l$ is the $L \times 1$ unit vector whose nonzero element is in the l-th position and $\otimes$ denotes the Kronecker product operator. Then for each sensor the discrete low-resolution observed image $\mathbf{g}_{l1,l2}$ can be written as

$$\mathbf{g}_{l1,l2} = \mathbf{D}_{l1,l2}\mathbf{H}_{l1,l2}\mathbf{f} + \mathbf{n}_{l1,l2} = \mathbf{W}_{l1,l2}\mathbf{f} + \mathbf{n}_{l1,l2}, \tag{4}$$

where $\mathbf{D}_{l1,l2} = \mathbf{D}_{l1} \otimes \mathbf{D}_{l2}$, denotes the $(N_1 \times N_2) \times (M_1 \times M_2)$ 2D downsampling matrix and $\mathbf{n}_{l1,l2}$ is modeled as independent white noise with variance $\beta_{l1,l2}^{-1}$.

We denote by $\mathbf{g}$ the sum of the upsampled low resolution images, that is,

$$\mathbf{g} = \sum_{u,v \in \mathcal{I}} \mathbf{D}_{u,v}^t \mathbf{g}_{u,v}. \quad (5)$$

3. DEGRADATION AND IMAGE MODELS

From the model in Eq. (4), we have that the probability density function of the $(l1, l2)$-th low resolution image $\mathbf{g}_{l1,l2}$ with $\mathbf{f}$ the 'true' high resolution image is given by

$$p(\mathbf{g}_{l1,l2}|\mathbf{f}, \beta_{l1,l2}) \propto \frac{1}{Z(\beta_{l1,l2})} \times \exp\left[-\frac{\beta_{l1,l2}}{2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2\right], \quad (6)$$

where $Z(\beta_{l1,l2}) = (2\pi/\beta_{l1,l2})^{(N_1 \times N_2)/2}$.

Taking into account that we have multiple low resolution images, the probability of $\mathbf{g}$ given $\mathbf{f}$ is

$$p(\mathbf{g}|\mathbf{f}, \underline{\beta}) = \prod_{(l1,l2) \in \mathcal{I}} p(\mathbf{g}_{l1,l2}|\mathbf{f}, \beta_{l1,l2}) \propto \frac{1}{Z_{noise}(\underline{\beta})} \times \exp\left[-\frac{1}{2} \sum_{(l1,l2) \in \mathcal{I}} \beta_{l1,l2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2\right], \quad (7)$$

where $\underline{\beta} = (\beta_{l1,l2}|(l1, l2) \in \mathcal{I})$, and $Z_{noise}(\underline{\beta}) = \prod_{(l1,l2) \in \mathcal{I}} Z(\beta_{l1,l2})$.

As prior model for $\mathbf{f}$ we use a simultaneous autoregression (SAR), that is [9]

$$p(\mathbf{f}|\alpha) = \frac{1}{Z_{prior}(\alpha)} \exp\{-\frac{1}{2}\alpha\, \mathbf{f}^t \mathbf{C}^t \mathbf{C}\, \mathbf{f}\}, \quad (8)$$

where the parameter α measures the smoothness of the 'true' image, $Z_{prior}(\alpha) = (\prod_{i,j} \lambda_{ij}^2)^{-1/2} (2\pi/\alpha)^{(M_1 \times M_2)/2}$ and $\lambda_{ij} = 1 - 2\phi(\cos(2\pi i/M_1) + \cos(2\pi j/M_2))$, $i = 1, 2, \ldots, M_1$, $j = 1, 2, \ldots, M_2$ and $\mathbf{C}$ is the Laplacian operator.

4. BAYESIAN ANALYSIS

The steps we follow in this paper to estimate the hyperparameters, α and $\underline{\beta}$, and the original image are

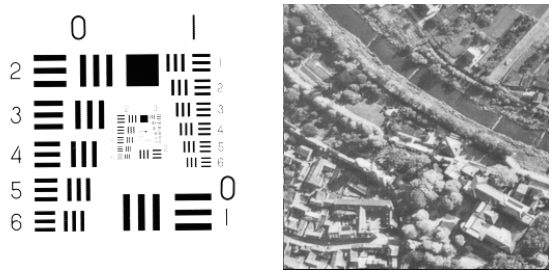
Step I: Estimation of the hyperparameters

$\hat{\alpha}$ and $\underline{\hat{\beta}} = (\hat{\beta}_{l1,l2}|(l1, l2) \in \mathcal{I})$ are first selected as

$$\hat{\alpha}, \underline{\hat{\beta}} = \arg\max_{\alpha, \underline{\beta}} \mathcal{L}_{\mathbf{g}}(\alpha, \underline{\beta}) = \arg\max_{\alpha, \underline{\beta}} \log p(\mathbf{g}|\alpha, \underline{\beta}), \quad (9)$$

where $p(\mathbf{g}|\alpha, \underline{\beta}) = \int_{\mathbf{f}} p(\mathbf{f}|\alpha) p(\mathbf{g}|\mathbf{f}, \underline{\beta}) d\mathbf{f}$.

Step II: Estimation of the original image



a b

Fig. 1. Original images: a) synthetic and b) real.

Once the hyperparameters have been estimated, the estimation of the original image, $\mathbf{f}_{(\hat{\alpha}, \underline{\hat{\beta}})}$, is selected as the image which minimizes

$$\hat{\alpha} \parallel \mathbf{C}\mathbf{f} \parallel^2 + \sum_{(l1,l2) \in \mathcal{I}} \hat{\beta}_{l1,l2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2, \quad (10)$$

resulting in

$$\mathbf{f}_{(\hat{\alpha}, \underline{\hat{\beta}})} = \mathbf{Q}\left(\hat{\alpha}, \underline{\hat{\beta}}\right)^{-1} \sum_{(l1,l2) \in \mathcal{I}} \hat{\beta}_{l1,l2} \mathbf{W}_{l1,l2}^t \mathbf{g}_{l1,l2}, \quad (11)$$

where $\mathbf{Q}(\hat{\alpha}, \underline{\hat{\beta}}) = \hat{\alpha}\mathbf{C}^t\mathbf{C} + \sum_{(l1,l2) \in \mathcal{I}} \hat{\beta}_{l1,l2} \mathbf{W}_{l1,l2}^t \mathbf{W}_{l1,l2}$.

It is important to note that the solution to Eq. (9) is obtained with the EM-algorithm with $\mathcal{X}^t = (\mathbf{f}^t, \mathbf{g}^t)$ and $\mathcal{Y} = \mathbf{g} = [\mathbf{0}\ \mathbf{I}]^{\mathbf{t}} \mathcal{X}$ and that the calculations involved in finding $\hat{\alpha}, \underline{\hat{\beta}}$ and $\mathbf{f}_{(\hat{\alpha}, \underline{\hat{\beta}})}$ can be performed using the general framework for frequency domain multi-channel signal processing developed in [5].

Let us now examine matrix $\mathbf{Q}(\alpha, \underline{\beta})$. We note that when fewer than $L \times L$ low resolution observations are available or when the shifts in those low resolution images do not satisfy the conditions in [2] and [3] this matrix can only be inverted and used in Eq. 11 because of the presence of $\mathbf{C}$. It is therefore important then to examine the role played by the prior model in Eq. (8) and also the accuracy of the estimated hyperparameters as a function of the number of low resolution observations, q. This is done experimentally, as described in detail in the next section.

5. EXPERIMENTAL RESULTS

A number of simulations have been performed with the proposed algorithm over a set of images to evaluate its performance depending on the number of available low resolution images.

The performance of the proposed algorithm was evaluated by measuring the peak signal-to-noise ratio (PSNR) defined as PSNR$= 10 \times \log_{10}[M_1 \times M_2 \times 255^2 / \parallel \mathbf{f} - \hat{\mathbf{f}} \parallel^2]$, where $\mathbf{f}$ and $\hat{\mathbf{f}}$ are the original and estimated high resolution images, respectively.

Results are presented on two different images; a synthetic image with numbers and patterns, depicted in Fig.1a, and a real image shown in Fig.1b.

Experiment I: According to Eq. (4) the synthetic high resolution image in Fig.1a was blurred using Eq. (2) obtaining $\mathbf{u} = \mathbf{H}\mathbf{f}$. Then the blurred high resolution image, $\mathbf{u}$, was downsampled with $L = 4$, thus obtaining a set of 16 low resolution images, $u_{l1,l2}(x, y) = u(L_1 x + l1, L_2 y + l2)$, $x, y = 0, \ldots, \frac{M_1}{L} - 1$, $l1, l2 = 0, \ldots, 3$. Gaussian noise was added to each low resolution image to obtain three sets of sixteen low resolution images,

Table 1. Noise variances for the synthetic low resolution image set with SNR of 40dB.

$\beta^{-1}_{l1,l2}$	0	1	2	3
0	0.40	0.40	0.39	0.40
1	0.39	0.40	0.39	0.40
2	0.38	0.39	0.39	0.38
3	0.39	0.41	0.39	0.40

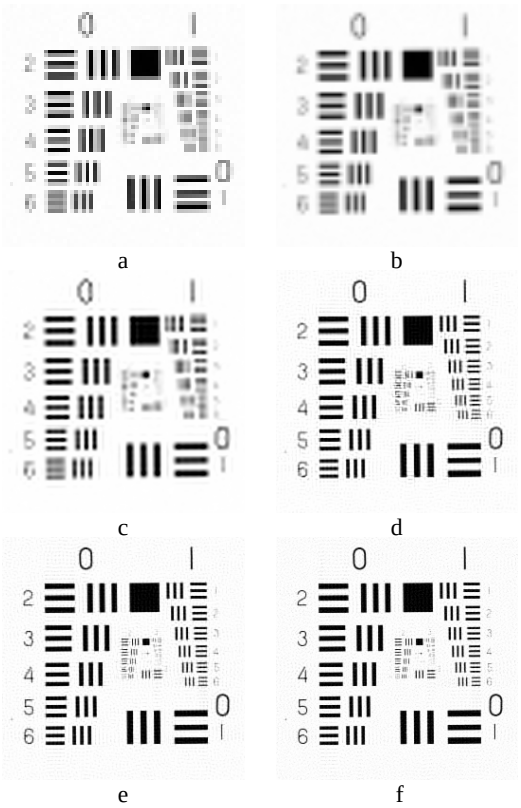


Fig. 2. (a) zero order hold, (b) bilinear interpolation, (c)-(f) results with the proposed method using 1, 4, 8 and 16 low resolution images.

$\mathbf{g}_{l1,l2}$, with 20, 30 and 40dB SNR. The noise variances for the 40dB set of images are shown in Table 1.

In order to test the performance of the proposed algorithm we ran it on different sets of q randomly chosen low resolution images with $1 \leq q \leq 16$. Figure 2a depicts the zero-order hold upsampled image of $\mathbf{g}_{0,0}$ for 40dB SNR (PSNR=14.07dB), bilinear interpolation of $\mathbf{g}_{0,0}$ is shown Fig. 2b (PSNR=15.62dB) and the estimated high-resolution images using 1, 4, 8 and 16 low resolution images are depicted in Fig. 2c–f, respectively. Estimated noise parameters, $\hat{\beta}$, using the proposed algorithm are shown in Table 2. Examining the table we conclude that the proposed method produces accurate estimations for all the low resolution image noise variances especially when the number of input images is high. Visual inspection shows that the proposed method outperforms zero order hold and bilinear interpolation for all the studied cases and, as expected, produces better results as the number of input images increases. PSNR evolution against the number of low resolution input images is shown in Fig. 3. Numerical results show that the

Table 2. Estimated noise variances for the synthetic low resolution image set with SNR of 40 dB.

1 image	$\hat{\beta}^{-1}_{0,0}$ 0.61				
4 images	$\hat{\beta}^{-1}_{0,0}$ 0.31	$\hat{\beta}^{-1}_{0,1}$ 1.47	$\hat{\beta}^{-1}_{2,2}$ 1.72	$\hat{\beta}^{-1}_{3,1}$ 4.02	
8 images	$\hat{\beta}^{-1}_{l1,l2}$	0	1	2	3
	0	0.86	0.92	N/A	0.60
	1	N/A	N/A	0.66	0.95
	2	N/A	N/A	0.72	N/A
	3	N/A	1.82	0.79	N/A
16 images	$\hat{\beta}^{-1}_{l1,l2}$	0	1	2	3
	0	0.38	0.37	0.36	0.35
	1	0.40	0.40	0.37	0.40
	2	0.36	0.35	0.35	0.38
	3	0.38	0.41	0.37	0.38

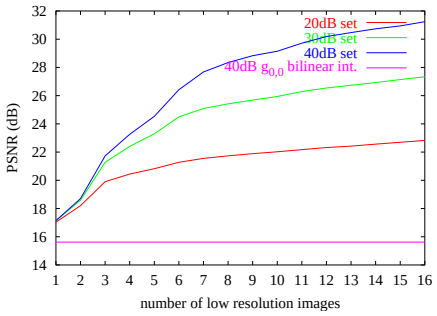


Fig. 3. PSNR evolution with the number of low resolution images for synthetic image.

proposed method provides a clear improvement even in the case when severe noise is present although higher improvements are obtained as the noise decreases. Note that the proposed algorithm always outperforms bilinear interpolation even when only one image is used (PSNR for the reconstructed image using just one low resolution input image is 17.14dB (see Fig. 2c), and it increases monotonically to 31.24dB with the number of images).

Experiment II: We also tested the proposed method on a real image. The original image in Fig.1b was blurred and downsampled as in the previous experiment and Gaussian noise was added to each low resolution image to obtain three sets of sixteen low resolution images with 10, 20 and 30dB SNR. Table 3 shows the noise variances for the 20dB set.

The algorithm was run on each low resolution set using from 1 to 16 randomly chosen low resolution images. Figure 4b–d de-

Table 3. Noise variances for the low resolution real image set with SNR of 20dB.

$\beta^{-1}_{l1,l2}$	0	1	2	3
0	24.55	24.61	24.62	24.74
1	24.64	24.54	24.68	24.62
2	24.59	24.69	24.56	24.76
3	24.68	24.70	24.62	24.60

Table 4. Estimated noise variances for the low resolution real image set with SNR of 20 dB.

1 image	$\hat{\beta}_{0,0}^{-1}$ 29.85				
5 images	$\hat{\beta}_{0,0}^{-1}$ 27.17	$\hat{\beta}_{1,0}^{-1}$ 31.07	$\hat{\beta}_{1,3}^{-1}$ 29.58	$\hat{\beta}_{2,2}^{-1}$ 31.39	$\hat{\beta}_{2,3}^{-1}$ 29.58
16 images	$\hat{\beta}_{l1,l2}^{-1}$	0	1	2	3
	0	25.25	25.30	26.21	28.23
	1	24.78	25.27	27.38	26.62
	2	26.25	26.35	26.24	25.83
	3	26.97	28.03	25.26	25.40

picts the resulting images for the 20dB SNR set using 1, 5 and 16 low resolution input images and the corresponding estimated noise parameters are presented in table 4. Bilinear interpolation result for $\mathbf{g}_{0,0}$, shown in Figure 4b with PSNR 20.61dB, is reported for comparison purposes. Again, the proposed method provides good reconstructions and accurate noise variance estimates even when only one image is used. Note that using only 5 low resolution images the results are almost indistinguishable from the best reconstruction (Figure 4f). This means that the prior model assist in accurately recovering the high resolution image even when we have little information and moderate to high noise. The method clearly improves the PSNR as the number of low resolution images increases, ass seen in Fig. 5. Note that most of the improvement is achieved with a small number of images. The proposed iterative algorithm typically needed 10 to 20 iterations to converge. Each iteration took a maximum of 15.5 seconds on a Pentium 4 1700.

6. CONCLUSIONS

A new method to estimate a high resolution image has been proposed. Using block-semi circulant matrices all the matrix calculations can be performed in the Fourier domain. The approach followed can be used with any number of low resolution images from 1 to L^2 since the prior model accurately recovers the high resolution image even in the case where just one or very few input images are provided. The proposed method has been validated experimentally.

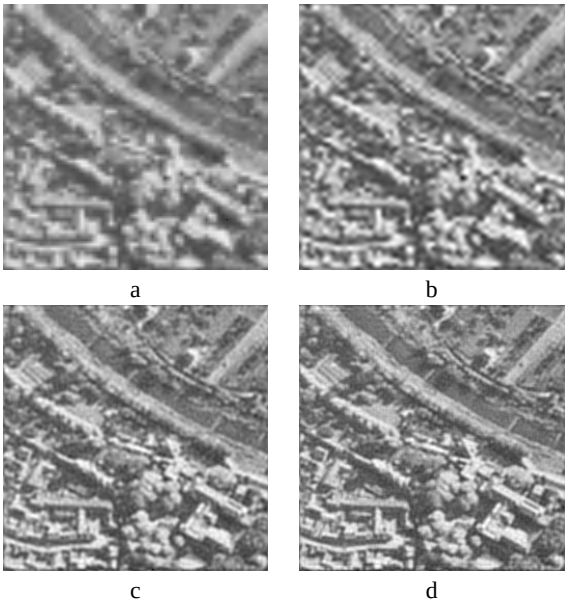


Fig. 4. (a) bilinear interpolation, (b)-(d) results with the proposed method using 1, 5 and 16 low resolution images.

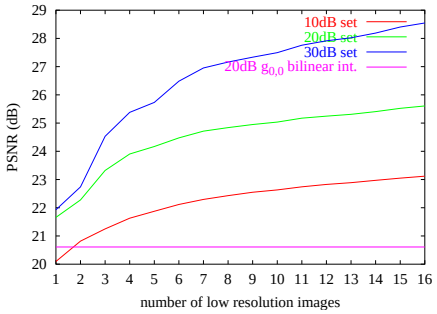


Fig. 5. PSNR evolution with the number of low resolution images for real image

7. REFERENCES

[1] S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences. A comprehensive review with directions for future research," Tech. Rep., Laboratory for Image and Signal Analysis, University of Notre Dame, 1998.

[2] S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 6, pp. 10131027, 1990.

[3] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. on Image Processing*, vol. 6, pp. 1646–1658, 1997.

[4] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, 2002.

[5] A. K. Katsaggelos, K. T. Lay, and N. P. Galatsanos, "A general framework for frequency domain multi-channel signal processing," *IEEE Trans. Image Processing,*, vol. 2, no. 3, pp. 417–420, 1993.

[6] N. K. Bose and K. J. Boo, "High-resolution image reconstruction with multisensors," *Int. Journ. Imaging Systems and Technology*, vol. 9, pp. 141–163, 1998.

[7] M. K. Ng and A. M. Yip, "A fast MAP algorithm for high-resolution image reconstruction with multisensors," *Multidimensional Systems and Signal Processing*, vol. 12, pp. 143–164, 2001.

[8] B. C. Tom, N. P. Galatsanos, and A. K. Katsaggelos, "Reconstruction of a high resolution image from multiple low resolution images," in *Super-Resolution Imaging*, S. Chaudhuri, Ed., chapter 4, pp. 73–105. Kluwer Academic Publishers, 2001.

[9] B. D. Ripley, *Spatial Statistics*, John Wiley, 1981.