PARAMETER ESTIMATION IN SUPER-RESOLUTION IMAGE RECONSTRUCTION PROBLEMS

Javier Abad[a], *Miguel Vega*[b], *Rafael Molina*[a] *and Aggelos K. Katsaggelos*[c*]

a) Dpto. de Ciencias de la Computación e I.A., Universidad de Granada, 18071 Granada, Spain
b) Dpto. de Lenguajes y Sistemas Informáticos, Universidad de Granada, 18071 Granada, Spain
c) Dept. of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208-3118
e-mail: abad@decsai.ugr.es, mvega@ugr.es, rms@decsai.ugr.es, aggk@ece.nwu.edu

ABSTRACT

In this paper we consider the estimation of the unknown hyperparameters for the problem of reconstructing a high-resolution image from multiple undersampled, shifted, degraded frames with subpixel displacement errors. We derive mathematical expressions for the iterative calculation of the maximum likelihood estimate (mle) of the unknown hyperparameters given the low resolution observed images. Experimental results are presented for evaluating the accuracy of the proposed method.

1. INTRODUCTION

High resolution images can be obtained directly from high precision optics and CCD devices. However, due to hardware and cost limitations, as well as limitations arising from the underlying physics of the imaging problem, imaging systems often provide us with only multiple low resolution images. Low resolution images are common in many imaging applications, such us remote sensing, surveillance and astronomy.

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] and [2] and references therein). However, as reported in [3], not much work has been devoted to the efficient calculation of the reconstruction or the estimation of the associated hyperparameters (see, however, [4], [5], and [6].

In this paper we use the general framework for frequency domain multi-channel signal processing developed in [7] and [8] to tackle the estimation of hyperparameters in high resolution problems. With the use of block-semi circulant (BSC) matrices all the matrix calculations involved in the hyperparameter mle can be performed in the Fourier domain. The proposed approach can be used to assign the same hyperparameter to all low resolution images or make them image dependent.

The rest of the paper is organized as follows. The problem formulation is described in section 2. Section 3 describes the high resolution prior image model and the process to obtain the low resolution images from the high resolution one. The application of the Bayesian paradigm to calculate the maximum *a posteriori* high resolution image and to 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 sensor array with $L_1 \times L_2$ sensors, each sensor having $N_1 \times N_2$ pixels and the size of each sensing element being $T_1 \times T_2$. Our aim is to reconstruct an $M_1 \times M_2$ high resolution image, where $M_1 = L_1 \times N_1$ and $M_2 = L_2 \times N_2$, from $L_1 \times L_2$ low-resolution observed images.

Note that in order for our goal to make sense we need to assume that the original high-resolution scene is bandlimited to wavenumbers $L_1/(2T_1)$ and $L_2/(2T_2)$ along the horizontal and vertical directions, respectively. To maintain the aspect ratio of the reconstructed image we consider the case where $L_1 = L_2 = L$.

In the ideal case, the low resolution sensors are shifted with respect to each other by a value proportional to $T_1/L \times T_2/L$ (note that if the sensors are shifted by values proportional to $T_1 \times T_2$ the high-resolution image reconstruction problem becomes singular). However, in practice there can be small perturbations around those ideal locations. Thus, for $l_1, l_2 = 0, \ldots, L-1$, the horizontal and vertical displacements $d^x_{l1,l2}$ and $d^y_{l1,l2}$ of the $[l_1, l_2]$-th sensor with respect to the $[0, 0]$-th reference sensor are given by

$$d^x_{l1,l2} = \frac{T_1}{L}(l_1 + \epsilon^x_{l1,l2}) \text{ and } d^y_{l1,l2} = \frac{T_2}{L}(l_2 + \epsilon^y_{l1,l2}), \qquad (1)$$

where $\epsilon^x_{l1,l2}$ and $\epsilon^y_{l1,l2}$ denote respectively the normalized horizontal and vertical displacement errors. We assume that $|\epsilon^x_{l1,l2}| < 1/2$ and $|\epsilon^y_{l1,l2}| < 1/2$ with $\epsilon^x_{0,0} = \epsilon^y_{0,0} = 0$. The normalized horizontal and vertical displacement are assumed to be known (see [9, 10] for details). In [11] we can find an approach where the displacements are assumed unknown and are estimated simultaneously with the high-resolution image.

3. IMAGE AND DEGRADATION MODELS

Let $\mathbf{f}$ be the $(M_1 \times M_2) \times 1$ high resolution image and $\mathbf{g}_{l1,l2}$ the $(N_1 \times N_2) \times 1$ observed low resolution image from the $(l1, l2)$-th sensor, $(l1, l2) \in \{0, \ldots L-1\}^2$. Our goal is to reconstruct $\mathbf{f}$ from $\left\{\mathbf{g}_{l1,l2} \mid (l1, l2) \in \{0, \ldots L-1\}^2\right\}$. In order to apply the Bayesian paradigm to this problem we define next our image and high to low degradation models.

3.1. Image Model

Our prior knowledge about the smoothness of the object luminosity distribution makes it possible to model the distribution of $\mathbf{f}$ by a simultaneous autoregression (SAR):

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

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

3.2. Model for obtaining the low-resolution observed 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. This image represents a blurred version of the original high-resolution one, according to

$$\mathbf{f}^{l1,l2} = \mathbf{H}_{l1,l2}\mathbf{f}, \quad (3)$$

where $\mathbf{H}_{l1,l2}$ is an $(M_1 \times M_2) \times (M_1 \times M_2)$ matrix and may have different forms. In [12] h_l^i has the form

$$h_l^i(u) = \begin{cases} \frac{1}{L} & u = -(L-1), \ldots, 0 \\ 0 & \text{otherwise} \end{cases} . \quad (4)$$

Note that in this case, $h_{l1}^1 = h_{l2}^2, \forall i, \epsilon_l^i = 0$, the normalized horizontal and vertical displacement errors in Eq. (1) satisfy $\epsilon_{l1,l2}^x = \epsilon_{l1,l2}^y = 0$ and $\mathbf{H}_{l1,l2} = \mathbf{H}, \forall l1, l2 = 0, \ldots, L-1$.

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

$$\mathbf{D}_{l1} = \mathbf{I}_{N_1} \otimes \mathbf{e}_l^t, \qquad \mathbf{D}_{l2} = \mathbf{I}_{N_2} \otimes \mathbf{e}_l^t, \quad (5)$$

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

$$\mathbf{g}_{l1,l2} = \mathbf{D}_{l1,l2}\mathbf{H}_{l1,l2}\mathbf{f} + \mathbf{v}_{l1,l2}, \quad (6)$$

where

$$\mathbf{D}_{l1,l2} = \mathbf{D}_{l1} \otimes \mathbf{D}_{l2}, \quad (7)$$

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

If $\mathbf{W}_{l1,l2}$ denotes the $(N_1 \times N_2) \times (M_1 \times M_2)$ matrix

$$\mathbf{W}_{l1,l2} = \mathbf{D}_{l1,l2}\mathbf{H}_{l1,l2}, \quad (8)$$

then we have

$$p(\mathbf{g}_{l1,l2}|\mathbf{f}, \beta_{l1,l2}) \propto \frac{1}{Z(\beta_{l1,l2})} \times \exp\left[-\frac{\beta_{l1,l2}}{2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2\right], \quad (9)$$

where $Z(\beta_{l1,l2}) = (2\pi/\beta_{l1,l2})^{(N_1 \times N_2)/2}$. We denote by $\mathbf{g}$ the sum of the upsampled low resolution images, that is,

$$\mathbf{g} = \sum_{u=0}^{L-1}\sum_{v=0}^{L-1} \mathbf{D}_{u,v}^t \mathbf{g}_{u,v}. \quad (10)$$

Then

$$p(\mathbf{g}|\mathbf{f}, \underline{\beta}) \propto \frac{1}{Z_{noise}(\underline{\beta})} \times \exp\left[-\frac{1}{2}\sum_{l1=0}^{L-1}\sum_{l2=0}^{L-1} \beta_{l1,l2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2\right], \quad (11)$$

where $\underline{\beta} = (\beta_{l1,l2} \mid (l1, l2) \in \{0, \ldots, L-1\}^2)$, and $Z_{noise}(\underline{\beta}) = \prod_{l1=0}^{L-1}\prod_{l2=0}^{L-1} Z(\beta_{l1,l2})$.

4. BAYESIAN ANALYSIS

In this paper we follow the steps described below to estimate the hyperparameters, α and $\underline{\beta}$, and the original image.

Step I: Estimation of the hyperparameters

$\hat{\alpha}$ and $\hat{\underline{\beta}} = (\hat{\beta}_{l1,l2}|(l1, l2) \in \{0, \ldots, L-1\}^2)$ are first selected as

$$\hat{\alpha}, \hat{\underline{\beta}} = \arg\max_{\alpha, \underline{\beta}} \mathcal{L}_{\mathbf{g}}(\alpha, \underline{\beta}) = \arg\max_{\alpha, \underline{\beta}} \log p(\mathbf{g}|\alpha, \underline{\beta}), \quad (12)$$

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

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

$$\hat{\alpha} \parallel \mathbf{C}\mathbf{f} \parallel^2 + \sum_{l1=0}^{L-1}\sum_{l2=0}^{L-1} \hat{\beta}_{l1,l2} \parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f} \parallel^2, \quad (13)$$

so, we have

$$\mathbf{f}_{(\hat{\alpha}, \hat{\underline{\beta}})} = \mathbf{Q}\left(\hat{\alpha}, \hat{\underline{\beta}}\right)^{-1} \sum_{l1=0}^{L-1}\sum_{l2=0}^{L-1} \hat{\beta}_{l1,l2} \mathbf{W}_{l1,l2}^t \mathbf{g}_{l1,l2}, \quad (14)$$

where

$$\mathbf{Q}(\hat{\alpha}, \hat{\underline{\beta}}) = \hat{\alpha}\mathbf{C}^t\mathbf{C} + \sum_{l1=0}^{L-1}\sum_{l2=0}^{L-1} \hat{\beta}_{l1,l2} \mathbf{W}_{l1,l2}^t \mathbf{W}_{l1,l2}. \quad (15)$$

Note that we are using maximum likelihood for estimating the hyperparameter and *maximum a posteriori* (MAP) for estimating $\mathbf{f}$. Furthermore, although steps I and II are separated, the iterative scheme proposed next performs both estimations simultaneously.

The estimation process we are using could be performed within the so called hierarchical Bayesian approach [13] by including hyperpriors on the unknown hyperparameter $\hat{\alpha}$ and hypervector $\hat{\underline{\beta}}$. However, the possibility of incorporating additional knowledge on them by means of gamma or other distributions will not be discussed here (see [13]).

Differentiating $-2\mathcal{L}_{\mathbf{g}}(\alpha, \underline{\beta})$ with respect to α and $\underline{\beta}$ so as to find the conditions satisfied at the maxima, we have

$$\parallel \mathbf{C}\mathbf{f}_{(\alpha, \underline{\beta})} \parallel^2 + \mathrm{tr}[\mathbf{Q}(\alpha, \underline{\beta})^{-1}\mathbf{C}^t\mathbf{C}] = \frac{M_1 \times M_2}{\alpha}, \quad (16)$$

$$\parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f}_{(\alpha, \underline{\beta})} \parallel^2 + \mathrm{tr}[\mathbf{Q}(\alpha, \underline{\beta})^{-1}\mathbf{W}_{l1,l2}^t\mathbf{W}_{l1,l2}] = \frac{N_1 \times N_2}{\beta_{l1,l2}}, \quad \text{for } l1, l2 = 0, \ldots, L-1. \quad (17)$$

We will use the following algorithm for the estimation of the hyperparameters and the high-resolution image

1. Choose α^0 and $\underline{\beta}^0$.
2. Compute $\mathbf{f}_{(\alpha^0,\underline{\beta}^0)}$ using Eq. (14) with $\hat{\alpha} = \alpha^0$, $\hat{\underline{\beta}} = \underline{\beta}^0$.
3. Repeat for $k = 1, 2, \ldots$
   i) Calculate α^k and $\underline{\beta}^k$ by substituting α^{k-1} and $\underline{\beta}^{k-1}$ in the left hand side of Eqs. (16) and (17).
   ii) Compute $\mathbf{f}_{(\alpha^k,\underline{\beta}^k)}$ by Eq. (14) with $\hat{\alpha} = \alpha^k$, $\hat{\underline{\beta}} = \underline{\beta}^k$.

   until $\parallel \mathbf{f}_{(\alpha^k,\underline{\beta}^k)} - \mathbf{f}_{(\alpha^{k-1},\underline{\beta}^{k-1})} \parallel$ is less than a prescribed bound.

Note that if the same hyperparameter is used for certain low resolution observations, Eqs. (16) and (17) become easier to calculate.

Equations (16) and (17) can also be obtained with the EM-algorithm [14] with $\mathcal{X}^t = (\mathbf{f}^t, \mathbf{g}^t)$ and $\mathcal{Y} = \mathbf{g} = [\mathbf{0}\ \ \mathbf{I}]^{\mathbf{t}}\mathcal{X}$ to iteratively increase $\mathcal{L}_{\mathbf{g}}(\alpha, \underline{\beta})$.

5. EXPERIMENTAL RESULTS

A number of simulations have been performed with the proposed algorithm over a set of images. Here we present results on one image. The criterion

$$\frac{\parallel \mathbf{f}_{(\alpha^k,\underline{\beta}^k)} - \mathbf{f}_{(\alpha^{k-1},\underline{\beta}^{k-1})} \parallel^2}{\parallel \mathbf{f}_{(\alpha^{k-1},\underline{\beta}^{k-1})} \parallel^2} < 10^{-6}$$

was used for terminating iterations. We set $\mathbf{f}^0 = \mathbf{g}$, where $\mathbf{g}$ has been defined in Eq. (10), and the initial values of the hyperparameters

$$\begin{aligned} \frac{1}{\alpha^0} &= \frac{\parallel \mathbf{C}\mathbf{f}^0 \parallel^2}{M_1 \times M_2}, \\ \frac{1}{\beta^0_{l1,l2}} &= \frac{\parallel \mathbf{g}_{l1,l2} - \mathbf{W}_{l1,l2}\mathbf{f}^0 \parallel^2}{N_1 \times N_2}, \end{aligned}$$

for $l1, l2 = 0, \ldots, L-1$.

The performance of the restoration algorithms was evaluated by measuring the SNR improvement, $\Delta_{SNR} = 10 \times \log_{10} [\parallel \mathbf{f} - \mathbf{g} \parallel^2 / \parallel \mathbf{f} - \hat{\mathbf{f}} \parallel^2]$, where $\mathbf{f}$ and $\hat{\mathbf{f}}$ are the original and estimated high resolution images, respectively.

According to Eq. (6) the original image, shown in Fig. 1a, was blurred using Eq. (4), to obtain $\mathbf{u} = \mathbf{H}\mathbf{f}$. Then $\mathbf{u}$ was downsampled with $L_1 = L_2 = 4$, thus obtaining 16 low resolution images, $u_{l1,l2}(x, y) = u(L_1x + l1, L_2y + l2)$, $x, y = 0, \ldots, \frac{M_1}{4} - 1$, $l1, l2 = 0, \ldots, 3$. Gaussian noise was added to each low resolution image to obtain three sets of sixteen low resolution images with 10, 20 and 30dB SNRs. The noise variances for each set of images are shown in table 1. Figure 1b depicts the zero-order hold upsampled image $\mathbf{g}_{00}$ for 30dB SNR. The initial image, $\mathbf{f}^0 = \mathbf{g}$, for the 30dB set according to Eq. (10) is shown in Fig. 1c and the estimated high-resolution one in Fig. 1d. Visual inspection shows that the proposed method provides a substantial improvement.

Numerical results for the three sets of images with SNRs of 10, 20 and 30dB are summarized in table 2. It is clear that the proposed method improves the SNR even in the case of severe noise

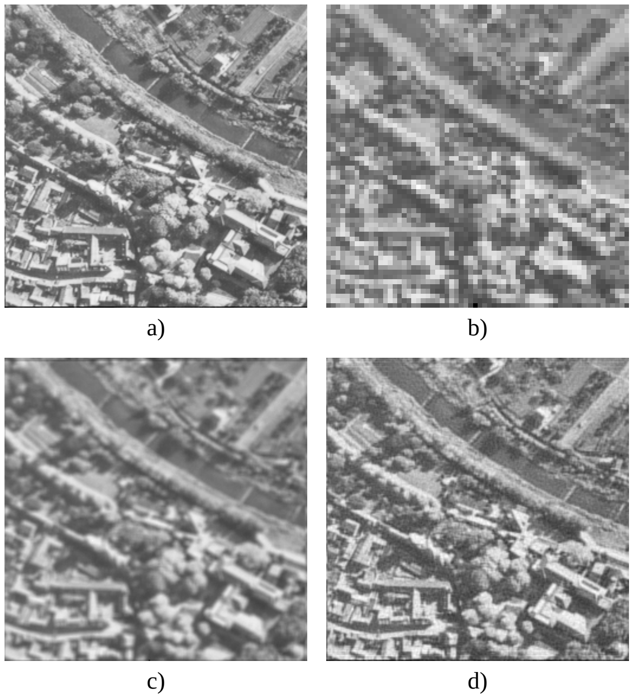

Fig. 1. a) Original 256×256 high-resolution image, b) zero-order hold for the 30dB SNR observation $\mathbf{g}_{00}$, c) initial high resolution image and d) estimated high resolution image.

although higher improvements are obtained as the noise decreases. Each low resolution set of 16 observed images was bilinearly interpolated to obtain a 256×256 image. For all cases, the best results using bilinear interpolation are only slightly better than the initial image. Table 2 also shows the number of iterations needed for the method to reach convergence according to the utilized criterion. For all sets of images no more than 13 iterations were needed. Each iteration took about 15.5 seconds on a Pentium IV 1700.

The estimated image model parameters for the 10, 20 and 30dB SNR sets of images were $\alpha^{-1} = 217.44, 192.22$ and 200.30, respectively. The estimated noise variances for the low resolution images are presented in table 3. We conclude that the proposed method produces accurate estimations of the low resolution image variances for all SNRs.

6. CONCLUSIONS

A new method to estimate the unknown hyperparameters in a high resolution image reconstruction problem has been proposed. Using BSC matrices all the matrix calculations involved in the hyperparameter mle can be performed in the Fourier domain. The approach followed can be used to assign the same hyperparameter to all low resolution image hyperparameters or to make them image dependent. The proposed method has been validated experimentally.

Table 1. Noise variances for the low resolution image set with SNR of 10dB, 20db and 30db.

	$\sigma^2_{l1,l2}$	0	1	2	3
10dB	0	163.19	161.62	162.53	161.63
	1	163.11	162.27	163.54	161.37
	2	163.19	163.45	161.66	161.34
	3	163.11	162.18	162.31	161.82
20dB	0	14.79	14.75	14.73	14.72
	1	14.89	14.70	14.85	14.72
	2	14.81	14.81	14.71	14.70
	3	14.84	14.74	14.85	14.69
30dB	0	1.47	1.46	1.46	1.46
	1	1.47	1.46	1.47	1.46
	2	1.47	1.47	1.46	1.46
	3	1.47	1.46	1.46	1.46

Table 2. Summary of results for the three different low resolution image sets.

Low resolution SNR (dB)	10	20	30
Bilin. interp. Δ_{SNR} (dB)	0.52	0.24	0.21
Proposed alg. Δ_{SNR} (dB)	6.168	7.924	10.723
Iterations	4	13	13

Table 3. Estimated noise variances for the 10dB, 20dB and 30dB SNR low resolution image sets.

	$\beta^{-1}_{l1,l2}$	0	1	2	3
10dB	0	166.35	160.12	158.04	160.00
	1	160.06	160.98	162.36	160.71
	2	157.46	160.58	162.98	156.90
	3	164.79	165.44	155.55	160.97
20dB	0	15.71	15.07	14.27	14.43
	1	15.13	14.84	14.49	14.72
	2	14.88	14.89	14.44	13.96
	3	15.53	15.52	13.92	14.35
30dB	0	1.53	1.48	1.38	1.40
	1	1.47	1.46	1.36	1.38
	2	1.46	1.41	1.39	1.35
	3	1.51	1.54	1.28	1.34

7. REFERENCES

[1] A. Averbuch and Y. Keller, "FFT based image registration," in *2002 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2002, vol. 4, pp. 3608–3611.

[2] H. Shekarforoush, M. Berthod, and J. Zerubia, "Subpixel image registration by estimating the polyphase decomposition of cross power spectrum," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 1996, pp. 532–537.

[3] 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.

[4] N. K. Bose, S. Lertrattanapanich, and J. Koo, "Advances in superresolution using L-curve," *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 433–436, vol 2., 2001.

[5] N. Nguyen, *Numerical Algorithms for superresolution*, Ph.D. thesis, Stanford University, 2001.

[6] N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Trans. on Image Processing*, vol. 10, no. 4, pp. 573–583, 2001.

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

[8] M. R. Banham, N. P. Galatsanos, H. L. Gonzalez, and A. K. Katsaggelos, "Multichannel restoration of single channel images using a wavelet-based subband decomposition," *IEEE Transactions on Image Processing*, vol. 3, no. 6, pp. 821–833, 1994.

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

[10] 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.

[11] 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.

[12] D. Rajan and S. Chaudhuri, "An MRF-based approach to generation of super-resolution images from blurred observations," *Journal of Mathematical Imaging and Vision*, vol. 16, pp. 5–153, 2002.

[13] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, A. K. Katsaggelos, and J. Mateos, "Hyperparameter estimation in image restoration problems with partially-known blurs," *Optical Engineering*, vol. 41, no. 8, pp. 1845–1854, 2002.

[14] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data," *Journal of the Royal Statistics Society B*, vol. 39, pp. 1–38, 1972.