

BAYESIAN SUPER-RESOLUTION OF TEXT IMAGE SEQUENCES FROM LOW RESOLUTION OBSERVATIONS.

Francisco J. Cortijo¹, Salvador Villena², Rafael Molina¹ and Aggelos Katsaggelos³

¹Dpto. de Ciencias de la Computación e I.A.
Universidad de Granada,
18071 Granada, Spain

²Dpto. de Lenguajes y Sistemas Informáticos
Universidad de Granada,
18071 Granada, Spain

³Dept. of Electrical and Computer Engineering
Northwestern University,
Evanston, IL 60208-3118

ABSTRACT

This paper deals with the problem of reconstructing high-resolution text images from an incomplete set of undersampled, blurred, and noisy images shifted with subpixel displacement. We derive mathematical expressions for the calculation of the maximum *a posteriori* estimate of the high resolution image and the estimation of the parameters involved in the model. The method is tested on real text images and car plates, examining the impact of blurring and the number of available low resolution images on the final estimate.

1. INTRODUCTION

High resolution images can, in some cases, be obtained directly from high precision optics and charge coupled devices (CCDs). However, due to hardware and cost limitations, imaging systems often provide us with only multiple low resolution images. In addition, there is a lower limit as to how small each CCD can be, due to the presence of shot noise [1] and the fact that the associated signal to noise ratio (SNR) is proportional to the size of the detector [2].

Text data reconstruction from low resolution images is an interesting problem encountered in many application areas. Current approaches to the problem can be found in [3] (the *hallucination* algorithm) and in the papers by Capel and Zisserman ([4], [5]) where the high resolution image is estimated using a Bayesian approach.

Mateos *et al.* ([6],[7]) proposed a new method to estimate a high resolution image from low resolution images when no blurring was present in the observation process [6] and when the low resolution images are also blurred [7], a case that frequently appears in Astronomy (see [8] for instance) and Remote Sensing. In this paper we apply this method to text data reconstruction (spanish car's licence plates and text data captures) and we demonstrate its performance when applied to this kind of data.

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.

2. PROBLEM FORMULATION

Consider a camera sensor with $N_1 \times N_2$ pixels and assume that we have a set of q shifted images, $1 \leq q \leq L \times L$. 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 a 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). In this paper we assume that the normalized horizontal and vertical displacements are known (see [9, 10] for details).

Let $\vec{g}_{l1,l2}$ be the $(N_1 \times N_2) \times 1$ observed low resolution image acquired by the $(l1, l2)$ sensor. Our goal is to reconstruct \vec{f} , the $(M_1 \times M_2) \times 1$ high resolution image, from a set of q low resolution images $\vec{g}_{l1,l2}$, 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 $(l1, l2)$ sensor, $\vec{g}_{l1,l2}$, from \vec{f} can be modeled as a two stages process as follows. In the first stage, the optical distortion in the observation process is represented by $\vec{B}_{l1,l2}$, the $(M_1 \times M_2) \times (M_1 \times M_2)$ point spread function defining a systematic blur, assumed to be known, due, for example, to motion or out of focus blurring, defects in the camera optics, etc. The second stage models the CCD pixel resolution. Let $\vec{H}_{l1,l2}$ be an $(M_1 \times M_2) \times (M_1 \times M_2)$ integrating matrix that represents the way a set of pixels in the high resolution image affects each low resolution pixel. In this paper we use an $\vec{H}_{l1,l2}$ representing a linear space-invariant blurring system with impulse response

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

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

$$\vec{D}_{l1} = \vec{I}_{N_1} \otimes \vec{e}_l^t, \quad \vec{D}_{l2} = \vec{I}_{N_2} \otimes \vec{e}_l^t, \quad (2)$$

where \vec{I}_{N_i} is the $N_i \times N_i$ identity matrix, \vec{e}_l^t 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 $\vec{g}_{l1,l2}$ can be written as

$$\vec{g}_{l1,l2} = \vec{D}_{l1,l2} \vec{H}_{l1,l2} \vec{B}_{l1,l2} \vec{f} + \vec{n}_{l1,l2} = \vec{W}_{l1,l2} \vec{B}_{l1,l2} \vec{f} + \vec{n}_{l1,l2}, \quad (3)$$

where $\vec{D}_{l1,l2} = \vec{D}_{l1} \otimes \vec{D}_{l2}$, denotes the $(N_1 \times N_2) \times (M_1 \times M_2)$ 2D downsampling matrix and $\vec{n}_{l1,l2}$ is modeled as independent white noise with variance $\beta_{l1,l2}^{-1}$. We denote by \vec{g} the sum of the upsampled low resolution images, that is,

$$\vec{g} = \sum_{u,v \in \mathcal{I}} \vec{D}_{u,v}^t \vec{g}_{u,v}. \quad (4)$$

3. DEGRADATION AND IMAGE MODELS

From Eq. 3, the probability density function of $\vec{g}_{l1,l2}$, the $(l1, l2)$ low resolution image, with \vec{f} the ‘true’ high resolution image, $p(\vec{g}_{l1,l2} | \vec{f}, \beta_{l1,l2})$, is proportional to

$$\frac{1}{Z(\beta_{l1,l2})} \exp \left[-\frac{\beta_{l1,l2}}{2} \|\vec{g}_{l1,l2} - \vec{W}_{l1,l2} \vec{B}_{l1,l2} \vec{f}\|^2 \right] \quad (5)$$

where $Z(\beta_{l1,l2}) = (2\pi/\beta_{l1,l2})^{(N_1 \times N_2)/2}$ and $\beta_{l1,l2}$ the inverse of the noise variance.

Since we have multiple low resolution images, $p(\vec{g} | \vec{f}, \underline{\beta}) = \prod_{(l1,l2) \in \mathcal{I}} p(\vec{g}_{l1,l2} | \vec{f}, \beta_{l1,l2})$ and it is proportional to

$$\frac{1}{Z_{noise}(\underline{\beta})} \exp \left[-\frac{1}{2} \sum_{(l1,l2) \in \mathcal{I}} \beta_{l1,l2} \|\vec{g}_{l1,l2} - \vec{W}_{l1,l2} \vec{B}_{l1,l2} \vec{f}\|^2 \right] \quad (6)$$

where

$$\underline{\beta} = (\beta_{l1,l2} | (l1, l2) \in \mathcal{I}),$$

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

As prior model for \vec{f} we use a simultaneous autoregression (SAR) [11], that is

$$p(\vec{f} | \alpha) = \frac{1}{Z_{prior}(\alpha)} \exp \left\{ -\frac{1}{2} \alpha \vec{f}^t \vec{C}^t \vec{C} \vec{f} \right\}, \quad (7)$$

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 $\lambda_{ij} = 1 - 2\phi(\cos(2\pi i/M_1) + \cos(2\pi j/M_2))$, $i = 1, 2, \dots, M_1$, $j = 1, 2, \dots, M_2$ and \vec{C} is the Laplacian operator.

4. BAYESIAN ANALYSIS

Having defined the degradation and image models, the Bayesian analysis is performed to estimate the hyperparameters, α and $\underline{\beta}$, and the high-resolution image. In this paper we use the following two steps:

Step I: Estimation of the hyperparameters

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

$$\hat{\alpha}, \hat{\underline{\beta}} = \arg \max_{\alpha, \underline{\beta}} \mathcal{L}_{\vec{g}}(\alpha, \underline{\beta}) = \arg \max_{\alpha, \underline{\beta}} \log p(\vec{g} | \alpha, \underline{\beta}), \quad (8)$$

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

The solution to this equation is obtained with the EM-algorithm with $\mathcal{X}^t = (\vec{f}^t, \vec{g}^t)$ and $\mathcal{Y} = \vec{g} = [\mathbf{0} \quad \vec{\mathbf{1}}]^t \mathcal{X}$

Step II: Estimation of the high-resolution image

Once the hyperparameters have been estimated, the estimation of the high-resolution image, $\vec{f}_{(\hat{\alpha}, \hat{\underline{\beta}})}$, is selected to minimize

$$\hat{\alpha} \|\vec{C} \vec{f}\|^2 + \sum_{(l1,l2) \in \mathcal{I}} \hat{\beta}_{l1,l2} \|\vec{g}_{l1,l2} - \vec{W}_{l1,l2} \vec{B}_{l1,l2} \vec{f}\|^2, \quad (9)$$

which results in

$$\vec{f}_{(\hat{\alpha}, \hat{\underline{\beta}})} = \vec{Q}(\hat{\alpha}, \hat{\underline{\beta}})^{-1} \sum_{(l1,l2) \in \mathcal{I}} \hat{\beta}_{l1,l2} \vec{B}_{l1,l2}^t \vec{W}_{l1,l2}^t \vec{g}_{l1,l2}, \quad (10)$$

where

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

Note that the prior model in Eq. 7 plays an important role in the estimation of the high-resolution image and the hyperparameters. If we examine the matrix $\vec{Q}(\alpha, \underline{\beta})$ in Eq. 10 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 [12] and [13] this matrix would not be invertible without the presence of \vec{C} . It is therefore important to examine the quality of the reconstruction 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. It is important to note that the calculations involved in finding $\hat{\alpha}$, $\hat{\underline{\beta}}$ and $\vec{f}_{(\hat{\alpha}, \hat{\underline{\beta}})}$ can be performed using the general framework developed in [14].

5. EXPERIMENTAL RESULTS

A number of experiments were performed with the proposed algorithm over different sets of images to evaluate its behavior as a function of the number of available low resolution images. We have experimented with two kinds of high resolution real images: an spanish car’s plate and a text image. In both cases we have also investigated the effect of blurring the high resolution images.

The first experiments were applied to Fig. 1a. That figure shows a car’s plate captured with a Sony CCD camera. The full resolution image was 1360×1020 pixels in size. We have considered only the subimage that contains the region of interest (128×64 pixels in size). This image was not blurred, and the integrating function in Eq. 1 was applied to the original image obtaining $\vec{u} = \vec{H} \vec{f}$. The high resolution image, \vec{u} , was downsampled with $L = 4$, thus obtaining a set of 16 low resolution images, $u_{l1,l2}(x, y) = u(L_1 x + l1, L_2 y + l2)$, $x, y = 0, \dots, \frac{M_1}{L} - 1$, $l1, l2 = 0, \dots, 3$. Gaussian noise was added to each low resolution image to obtain a set of sixteen low resolution images, $\vec{g}_{l1,l2}$, with 40dB SNR. Figure 1b depicts the zero-order hold upsampled image of $\vec{g}_{0,0}$ for 40dB SNR.

We have ran the reconstruction algorithm on different sets of q randomly chosen low resolution images with $1 \leq q \leq 16$. Figures 1c to 1i depict the estimated high-resolution images using 1, 2, 4, 5, 8, 12 and 16 low resolution images, respectively. Visual inspection of the resulting images shows that the proposed method clearly outperforms zero-order hold upsample even when a few low resolution input images are used (for example, Fig. 1f shows a high resolution image calculated from 5 low resolution images).

Note also that the quality of the high resolution estimated image increases with the number of low resolution images (see Fig. 5) but improvement is not significant from 8 to 16 low resolution images. The visual quality of the estimated high resolution images obtained using 8 and 16 low resolution input images (Fig. 1g and h, respectively) are almost indistinguishable, which means that the prior model assists in accurately recovering the high resolution image even when little information is available.

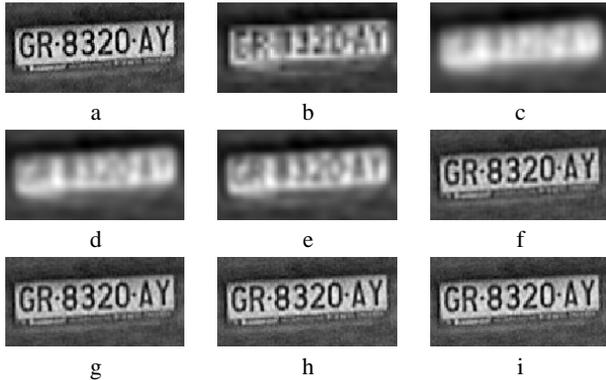


Figure 1: Spanish car's license plate capture: (a) high resolution original image (b) upsampled low resolution image (c)-(i) reconstruction using 1, 2, 4, 5, 8, 12 and 16 low resolution images.

We also investigated the performance of the algorithm when the high resolution image is blurred. Using Eq. 3 the high resolution image was blurred using a Gaussian blur with radius 5 and $\sigma = 5$. Figure 2a depicts the high resolution -blurred- image. Then, the integrating function in Eq. 1 was applied to the blurred image obtaining $\vec{u} = \vec{H}\vec{B}\vec{f}$. This high resolution image, \vec{u} , was downsampled with $L = 4$, thus obtaining a set of 16 low resolution images, $u_{l1,l2}(x, y) = u(L_1x + l1, L_2y + l2)$, $x, y = 0, \dots, \frac{M_1}{L} - 1, l1, l2 = 0, \dots, 3$. Gaussian noise was added to each low resolution image to obtain a set of sixteen low resolution images, $\vec{g}_{l1,l2}$, with 40dB SNR. Figure 2b depicts the zero-order hold upsampled image of $\vec{g}_{0,0}$ for 40dB SNR.

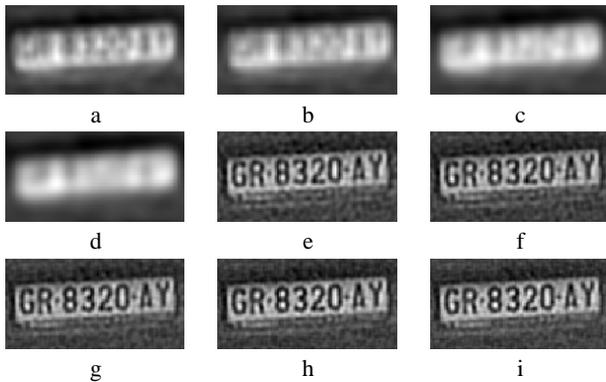


Figure 2: Blurred spanish car's license plate: (a) high resolution blurred image (b) upsampled low resolution image (c)-(i) reconstruction using 1, 2, 4, 5, 8, 12 and 16 low resolution images.

Once again, we ran the reconstruction algorithm on different sets of randomly chosen low resolution images. Figures 2c to 2i depict the estimated high-resolution images using 1, 2, 4, 5, 8, 12 and 16 low resolution images, respectively (the same numbers as in Fig. 1). Visual inspection shows again that the estimated high resolution images are much better than the zero-order hold upsamples even when a few low resolution input images are used. The quality of the high resolution estimated image also increases with the number of low resolution images (see Fig. 5).

We have also tested this method to improve low resolution text images. In Fig 3a we show a high resolution image (256×64 pixels in size). Initially, we did not blur this image and we obtained a set of 16 low resolution images (adding Gaussian noise with 40dB SNR) in the manner we have already described. Then we ran the reconstruction algorithm on different sets of randomly chosen low resolution images.

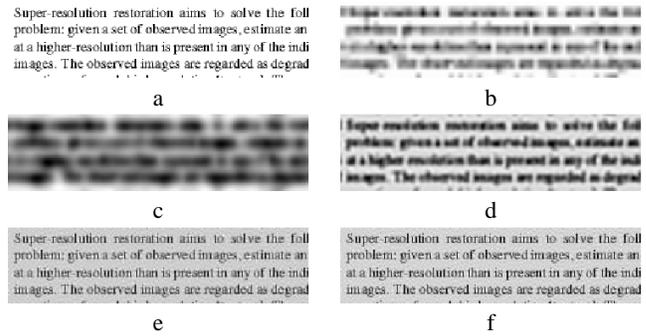


Figure 3: Text data capture (PDF viewer capture): (a) high resolution image (b) upsampled low resolution image (c)-(f) reconstruction using 1, 8, 12 and 16 low resolution images.

Figure 3b depicts the zero-order hold upsampled image of $\vec{g}_{0,0}$ and Figs. 3c-f depict the estimated high-resolution images using 1, 8, 12 and 16 low resolution images. Visual inspection of the resulting images shows that the high resolution estimated images are considerably better than zero-order hold upsamples (Fig. 1b). We need, however, a greater number of low resolution input images than in the previous experiments to obtain a high resolution image with good visual quality. In any case, the improvement is not significant with the number of low resolution images from 12 to 16 (see Fig. 5).

Finally, the high resolution image shown in Fig. 3a was blurred using a Gaussian blur with radius 2 and $\sigma = 2$. Figure 4a depicts the high resolution -blurred- image. The integrating function in Eq. 1 was applied to the blurred image, downsampled with $L = 4$ and Gaussian noise was added to each low resolution image. We obtained a set of sixteen low resolution images with 40dB SNR. Figure 4b depicts the zero-order hold upsampled image of $\vec{g}_{0,0}$ and Figs. 4c-f depict the estimated high-resolution images using 1, 8, 12 and 16 low resolution images, respectively (the same numbers as in Fig. 3).

Visual inspection of the resulting images shows that the estimated high resolution images are considerably better than the zero-order hold upsamples and that the quality of the high resolution estimated image also increases with the number of low resolution images (see Fig. 5). The improvement, however, is not as high as when the high resolution image is not blurred.

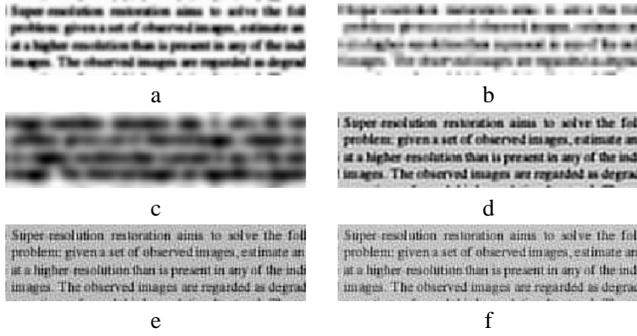


Figure 4: Blurred text data capture: (a) high resolution blurred image (b) upsampled low resolution image (c)-(f) reconstruction using 1, 8, 12 and 16 low resolution images.

Results are summarized in Fig.5. The performance of the proposed algorithm was evaluated by measuring the peak signal-to-noise ratio (PSNR) defined as $PSNR = 10 \times \log_{10} [M_1 \times M_2 \times 255^2 / \|\vec{f} - \vec{f}_{(\hat{\alpha}, \hat{\beta})}\|^2]$, where \vec{f} and $\vec{f}_{(\hat{\alpha}, \hat{\beta})}$ are the original and estimated high resolution images, respectively. PSNR evolution against the number of low resolution input images is shown in Fig. 5 for 40 dB low resolution images. Numerical results show that the proposed method provides a clear improvement when severe noise is present and most of the improvement is achieved when a low number of input images is used. This makes clear the importance of the prior model in the information recovering process. This method works well even when the image to be reconstructed is blurred. In this case, the improvement obtained with the number of low resolution input images is not as significant as when the original image is not blurred. We speculate that the number of required low resolution images to estimate a good high resolution image is inversely proportional to the scale of the objects present in the image (just compare *text* and *car's plate* images).

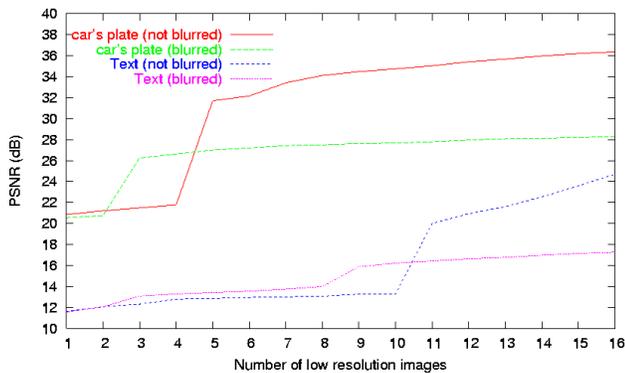


Figure 5: PSNR evolution with the no. of low resolution images

6. CONCLUSIONS

A new method to estimate a high resolution image from an incomplete set of blurred, undersampled low resolution images has been

proposed. The approach followed can be used with any number of low resolution images, 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 with text image sequences. We have shown that the number of low resolution images depends on the scale of the objects present in the image.

7. REFERENCES

- [1] Aizawa, K., Komatsu, T., Saito, T.: A scheme for acquiring very high resolution images using multiple cameras. In: IEEE Conference on Audio, Speech and Signal Processing. Volume 3. (1992) 289–292
- [2] Stark, H., Oskoui, P.: High-resolution image recovery from image-plane arrays, using convex projections. *Journal of the Optical Society A* **6** (1989) 1715–1726
- [3] Baker, S., Kanade, T.: Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24** (2002) 1167–1183
- [4] Capel, D., Zisserman, A.: Super-resolution enhancement of text image sequences. In: Proc. International Conf. on Pattern Recognition (ICPR). Volume I. (2000) 600–605
- [5] Capel, D., Zisserman, A.: Super-resolution from multiple views using learnt image models. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). Volume 2. (2001) 627–634
- [6] Mateos, J., Molina, R., Katsaggelos, A.K.: Bayesian high resolution image reconstruction with incomplete multisensor low resolution systems. To appear in Proc. International Conference on Acoustic, Speech and Signal Processing (2003)
- [7] J.Mateos, Vega, M., Molina, R., Katsaggelos, A.K.: Bayesian image estimation from an incomplete set of blurred, undersampled low resolution images. Submitted to 1st Iberian Conference on Pattern Recognition and Image Analysis (IBPRIA) (2003)
- [8] Molina, R., Nez, J., Cortijo, F., Mateos, J.: Image restoration in Astronomy. A Bayesian review. *IEEE Signal Processing Magazine* **18** (2001) 11–29
- [9] Bose, N.K., Boo, K.J.: High-resolution image reconstruction with multisensors. *Int. Journ. Imaging Systems and Technology* **9** (1998) 141–163
- [10] Ng, M.K., Yip, A.M.: A fast MAP algorithm for high-resolution image reconstruction with multisensors. *Multi-dimensional Systems and Sig. Proc.* **12** (2001) 143–164
- [11] Ripley, B.D.: *Spatial Statistics*. John Wiley (1981)
- [12] Kim, S.P., Bose, N.K., Valenzuela, H.M.: Recursive reconstruction of high resolution image from noisy undersampled multiframe. *IEEE Transactions on Acoustics, Speech and Signal Processing* **38** (1990) 1013–1027
- [13] Elad, M., Feuer, A.: Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans. on Image Processing* **6** (1997) 1646–1658
- [14] Katsaggelos, A.K., Lay, K.T., Galatsanos, N.P.: A general framework for frequency domain multi-channel signal processing. *IEEE Trans. Image Processing*, **2** (1993) 417–420