Journal of Mathematical Imaging and Vision 16: 5–15, 2002
© 2002 Kluwer Academic Publishers. *Manufactured in The Netherlands*.

An MRF-Based Approach to Generation of Super-Resolution Images from Blurred Observations

DEEPU RAJAN*
School of Biomedical Engineering, Indian Institute of Technology-Bombay, Powai, Mumbai 400076, India
deepu@ee.iitb.ac.in

SUBHASIS CHAUDHURI
Department of Electrical Engineering, Indian Institute of Technology-Bombay, Powai, Mumbai 400076, India
sc@ee.iitb.ac.in

Abstract. This paper presents a new technique for generating a high resolution image from a blurred image sequence; this is also referred to as super-resolution restoration of images. The image sequence consists of decimated, blurred and noisy versions of the high resolution image. The high resolution image is modeled as a Markov random field (MRF) and a maximum a posteriori (MAP) estimation technique is used for super-resolution restoration. Unlike other super-resolution imaging methods, the proposed technique does not require sub-pixel registration of given observations. A simple gradient descent method is used to optimize the functional. The discontinuities in the intensity process can be preserved by introducing suitable line processes. Superiority of this technique to standard methods of image expansion like pixel replication and spline interpolation is illustrated.

Keywords: super-resolution, image restoration, Markov random field

1. Introduction

The physical limitations of currently available image sensors impose a limit on the spatial resolution of images and videos. Such physical attributes include the size and density of detectors that form the sensor. In addition, the bandwidth limit set by the sampling rate also indirectly determines the resolution. Hence, low resolution images occur due to a combination of the effects of blurring due to the sensor point spread function (PSF) and aliasing as a consequence of undersampling. High spatial-resolution images are required in a variety of applications like remote sensing, aerial surveillance, medical imaging, high definition television and multimedia imaging. Super-resolution refers to the process of reconstructing a high resolution image from a single or multiple observations at a lower resolution. Such a restoration process entails upsampling the image (therefore, extrapolating the frequency spectrum), deblurring and undoing errors due to aliasing. The underlying philosophy of this method is to acquire more samples of the scene so as to get some additional information which can be utilized, while merging the samples to get a high resolution image. These samples can be acquired by sub-pixel shifts, by changing scene illumination or, as we propose in this paper, by changing the camera focus.

Tsai and Huang [29] were the first to propose a frequency domain approach to reconstruction of a high resolution image from a sequence of undersampled low resolution, noise-free images. Given a sufficient number of frames, the unaliased image is recovered by solving a set of equations in the frequency domain obtained from each of the aliased observations. However, this method requires a minimum number of low-resolution images, which may not be always available. Kim et al. [12] discuss a recursive least-squares

*On leave from the Department of Electronics, Cochin University of Science and Technology, Cochin 682022, India.

algorithm, also in the frequency domain, for the restoration of super-resolution images from noisy images. They extend the case to blurred observations in [13]. Ur and Gross [30] use the Papoulis-Brown generalized sampling theorem [16] to obtain an improved resolution picture from an ensemble of spatially shifted pictures. In all the aforementioned references, shifts are assumed to be known by the authors. Estimation of such shifts is the most difficult problem in super-resolution imaging and should be ideally avoided, if possible. A minimum mean squared error approach (MMSE) for mulitple image restoration, followed by interpolation of the restored images into a single high resolution image is presented in [26]. Tom and Katsaggelos [28] combine the sub-problems of registration, restoration and interpolation and formulate the problem using the maximum likelihood (ML) approach; the objective function is minimized with respect to noise variances and the blur. But since the objective function is too complex, the authors resort to the expectation-maximization (EM) algorithm to simplify it. Further, a quasi-Newton minimization technique is performed in the M-step of the EM algorithm for determining the sub-pixel shifts. Our contention, as stated earlier, is that such complex minimization procedures for intractable objective functions can be avoided by using low resolution images devoid of any shift. An iterative back-projection method is used in [10], wherein a guess of the high resolution output image is updated according to the error between the observed and the low resolution images obtained by simulating the imaging process. But back-projection methods can be used only for those blurring processes for which such an operator can be calculated. A similar approach is also described in [11]. Projection onto convex sets (POCS)-based methods are described in [17, 27]. Although these methods are quite simple, the interpretation of the result in terms of its spectral content is very difficult. Bascle et al. [1] present a method for simultaneous motion deblurring, focus deblurring and super-resolution from image sequences. However, the blurs have been assumed to be known. Chiang and Boult [5] use edge models and a local blur estimate to develop an edge-based super-resolution algorithm. But, in images with a large number of high frequency content, it is difficult to differentiate between significant and insignificant edges. Recently, Rajan and Chaudhuri [19] proposed a shape-from-shading-based super-resolution algorithm where bicubic spline interpolation was carried out on surface normals and albedo recovered from an image sequence and consequently, super-resolved images were reconstructed.

Several researchers have also addressed the above problem in a statistical framework. Shekarforoush et al. [25] use MRFs to model the images and obtain 3D high resolution visual information (albedo and depth) from a sequence of displaced low resolution images. In [24], an n dimensional extension of Papoulis's generalized sampling theorem is used to develop an iterative algorithm for 3D reconstruction of a Lambertian surface at a sub-pixel accuracy. Here too, the effect of sampling a scene at a higher rate is acquired by having interframe sub-pixel displacements. A MAP estimate of a super resolved image with Huber-MRF prior is described by Schultz and Stevenson in [22]. But they do not consider the case of blurred observations. They extended the method for extracting high resolution frames from video sequences [23]. After each frame is expanded using the algorithm in [22], motion is estimated using a block matching technique on individual frames. However, *a priori* knowledge about motion cannot be included in block matching techniques. Elad and Feuer [7] propose a unified methodology for super-resolution restoration from several geometrically warped, blurred, noisy and downsampled observations by combining ML, maximum a posteriori (MAP) and POCS approaches. This unified method generalizes these approaches only to a certain extent; MAP estimators with non-quadratic priors and POCS with L_∞ norm are not special cases of the model. Cheeseman et al. [4] describe another Bayesian approach for constructing a super-resolved surface by combining information from a set of images of the given surface. However, their model includes registration parameters, the point spread function (PSF) and camera parameters that are estimated first and subsequently, the surface reconstruction is carried out. Hardie et al. describe a joint MAP estimator for the high-resolution image and registration parameters to obtain sub-pixel translation estimates [20]. In [21], this method is extended to rotated and translated frames applied to infrared imaging systems.

In this paper, we present a new technique wherein an ensemble of decimated, blurred and noisy versions of an ideal high resolution image are used to generate a super-resolved image. This is an ill-posed inverse problem [9]. There is no spatial shifts between the observations but the images are captured with different camera blurs. The super-resolved image is modeled as an MRF and a MAP estimation technique is used. Since

there is no relative displacement between the input images, the need for estimating sub-pixel shifts does not arise. Also, the input images need not be focused as the algorithm carries out simultaneous restoration (deblurring) in the course of generating the super-resolution image. The proposed method is fast as a result of using a simple gradient-descent minimization of a convex cost function. One can obtain an improved result for super-resolution, if appropriate line fields [8] are included in the cost function. However, the computational requirement goes up. In situations where the computational cost is not an issue, simulated annealing can be used for optimization; a relatively faster approach would be to use the graduated non-convexity (GNC) algorithm [2]. Results for both these approaches are also presented to illustrate the efficacy of the proposed method.

In the next section we describe how low resolution images are generated from a high resolution image. In Section 3, we cast the super-resolution problem in a restoration framework. The cost function obtained using the MAP estimator is derived in this section. Section 4 presents experimental results, and conclusions are given in Section 5.

2. Low Resolution Image Formation

Suppose the low resolution image sensor plane is divided into $M_1 \times M_2$ square sensor elements and $Y = \{y_{i,j}\}$, $i = 0, \ldots, M_1 - 1$ and $j = 0, \ldots, M_2 - 1$, are the low resolution intensity values. If the downsampling parameters are q_1 and q_2 in the horizontal and vertical directions, respectively, then the high resolution image will be of size $q_1 M_1 \times q_2 M_2$. For notational ease, we assume $q_1 = q_2 = q$, and therefore the desired high-resolution image $\mathbf{z}$ will have intensity values $\{z_{k,l}\}$, $k = 0, \ldots, qM_1 - 1$ and $l = 0, \ldots, qM_2 - 1$. The forward process of obtaining $\{y_{i,j}\}$ from $\{z_{k,l}\}$ is written as

$$y_{i,j} = \frac{1}{q^2} \sum_{k=qi}^{(q+1)i-1} \sum_{l=qj}^{(q+1)j-1} z_{k,l} \tag{1}$$

i.e., the low resolution intensity is the average of the high resolution intensities over a neighborhood of q^2 pixels. This decimation model simulates the integration of light intensity that falls on the high-resolution detector.

Each of the decimated images is blurred by a different, but known linear space invariant blurring kernel. Elad and Feuer [6, 7] have shown that in this case super-resolution restoration is possible even if there is no relative motion between the input images. They derive the following necessary condition for super-resolution to be possible for images not represented parametrically:

$$q^2 \leq \min\{[2m + 1]^2 - 2, p\} \tag{2}$$

where $(2m + 1) \times (2m + 1)$ is the size of the blurring kernel and p is the number of input images. Hence, although more number of blurred observations of a scene do not provide any additional information in the same sense as sub-pixel shifts of the camera or changing illuminant directions do, it is, nevertheless, possible to achieve super-resolution with these blurred samples, provided Eq. (2) is satisfied. Even if only the relations among the blurring functions are known, as is the case in, say, depth from defocus problems [3, 18], it is tantamount to knowing all the blurs provided any one of them is known. Finally, i.i.d. zero mean Gaussian noise is added to the decimated and blurred images. Noise is assumed to be uncorrelated in different low resolution images.

Next, we formally state the problem by casting it in a restoration framework. There are p observed images $\{Y_i\}_{i=1}^{p}$ each of size $M_1 \times M_2$. These images are decimated, blurred and noisy versions of a single high resolution image Z of size $N_1 \times N_2$, where $N_1 = qM_1$ and $N_2 = qM_2$. If $\underline{\mathbf{y_i}}$ is the $M_1 M_2 \times 1$ lexicographically ordered vector containing pixels from the low resolution image Y_i, then a vector $\underline{\mathbf{z}}$ of size $q^2 M_1 M_2 \times 1$ containing pixels of the high resolution image can be formed by placing each of the $q \times q$ pixel neighborhoods sequentially so as to maintain the relationship between a low resolution pixel and its corresponding high resolution pixel. After incorporating the blur matrix and noise vector, the image formation model is written as

$$\underline{\mathbf{y_i}} = H_i D \underline{\mathbf{z}} + \underline{\mathbf{n_i}}, \quad i = 1, \ldots, p \tag{3}$$

where D is the decimation matrix of size $M_1 M_2 \times q^2 M_1 M_2$, H is the blurring matrix (PSF) of size $M_1 M_2 \times M_1 M_2$, $\underline{\mathbf{n_i}}$ is the $M_1 M_2 \times 1$ noise vector and p is the number of low resolution observations. The decimation matrix D consists of q^2 values of $\frac{1}{q^2}$ in each

row and has the form [22]

$$D = \frac{1}{q^2}\begin{bmatrix} 1\,1\ldots1 & & 0 \\ & 1\,1\ldots1 & \\ & \ddots & \\ 0 & & 1\,1\ldots1 \end{bmatrix} \quad (4)$$

Thus, the model indicates a collection of low resolution images, each of which differ from the others in the blur matrix, which is akin to changing the focus of a stationary camera looking at a stationary scene. Since we have assumed noise to be zero mean i.i.d, the multivariate pdf of $\underline{\mathbf{n_i}}$ is given by

$$P(\underline{\mathbf{n_i}}) = \frac{1}{(2\pi)^{\frac{M_1M_2}{2}}\sigma_\eta^{M_1M_2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\underline{\mathbf{n_i}}^T\underline{\mathbf{n_i}}\right\}, \quad (5)$$

where σ_η^2 denotes the variance of the noise process. Our problem now reduces to estimating $\underline{\mathbf{z}}$ given $\underline{\mathbf{y_i}}$'s, which is clearly an ill-posed, inverse problem.

3. Super-Resolution Using a MAP Estimator

The maximum a posteriori (MAP) estimation technique is used to obtain the high resolution image $\underline{\mathbf{z}}$ given the ensemble of low resolution image, i.e.,

$$\hat{\underline{\mathbf{z}}} = \arg\max_{\underline{\mathbf{z}}} P(\underline{\mathbf{z}} \mid \underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p) \quad (6)$$

From Bayes' rule, this can be written as

$$\hat{\underline{\mathbf{z}}} = \arg\max_{\underline{\mathbf{z}}} \frac{P(\underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p \mid \underline{\mathbf{z}})P(\underline{\mathbf{z}})}{P(\underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p)}. \quad (7)$$

Since the denominator is not a function of $\hat{\underline{\mathbf{z}}}$, equation (7) can be written as

$$\hat{\underline{\mathbf{z}}} = \arg\max_{\underline{\mathbf{z}}} P(\underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p \mid \underline{\mathbf{z}})P(\underline{\mathbf{z}}). \quad (8)$$

Taking the log of posterior probability,

$$\hat{\underline{\mathbf{z}}} = \arg\max_{\underline{\mathbf{z}}}[\log P(\underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p \mid \underline{\mathbf{z}}) + \log P(\underline{\mathbf{z}})]. \quad (9)$$

Hence, we need to specify the prior image density P($\underline{\mathbf{z}}$) and the conditional density $P(\underline{\mathbf{y}}_1, \underline{\mathbf{y}}_2, \ldots, \underline{\mathbf{y}}_p \mid \underline{\mathbf{z}})$.

3.1. Prior Model for $\underline{\mathbf{z}}$

MRF models have been widely used to solve vision problems because of their ability to model context dependency, since interpretation of visual information necessitates an efficient description of contextual constraints. The utility of MRF models arises from the Hammersley-Clifford theorem which describes the equivalence of the local property that characterizes an MRF and the global property which characterizes a Gibbs Random Field (GRF) [8, 15]. The high resolution image $\underline{\mathbf{z}}$ satisfying the Gibbs density function is now written as

$$P(\underline{\mathbf{z}}) = \frac{1}{Z}\exp\left\{-\sum_{c\in\mathcal{C}} V_c(\underline{\mathbf{z}})\right\} \quad (10)$$

where Z is a normalizing constant known as the partition function, $V_c(\cdot)$ is the clique potential and $\mathcal{C}$ is the set of all cliques in the image. In order to employ a simple and fast minimization technique like gradient descent, it is desirable to have a convex energy function. More importantly, the minimization procedure should not get trapped in a local minima. To this end, we consider pair wise cliques on a first order neighborhood and impose a quadratic cost which is a function of finite difference approximations of the first order derivative at each pixel location, i.e.,

$$V_c(\underline{\mathbf{z}}) = \frac{1}{\lambda}\sum_{k=1}^{N_1}\sum_{l=1}^{N_2}[(z_{k,l} - z_{k,l-1})^2 + (z_{k,l} - z_{k-1,l})^2] \quad (11)$$

where λ can be viewed as a "tuning" parameter. It can be interpreted as the penalty for departure from smoothness in $\underline{\mathbf{z}}$.

It is well known that in images, points having significant change in the image irradiance carry important information. In order to incorporate provisions for detecting such discontinuities, Geman and Geman [8] introduced the concept of line fields located on a dual lattice. We describe a prior using horizontal and vertical line fields in Section 3.4 and use GNC to optimize the corresponding cost function. As mentioned in the introduction, where computational issues do not arise, one could go in for simulated annealing (SA); we observe a significant improvement in the performance of the restoration process.

3.2. MAP Solution

From Eq. (9), since $\underline{\mathbf{n}}_i$'s are independent,

$$\hat{\underline{\mathbf{z}}} = \arg\max_{\underline{\mathbf{z}}} \left[\log \prod_{i=1}^{p} P(\underline{\mathbf{y}}_i \mid \underline{\mathbf{z}}) + \log P(\underline{\mathbf{z}}) \right] = \arg\max_{\underline{\mathbf{z}}} \left[\sum_{i=1}^{p} \log P(\underline{\mathbf{y}}_i \mid \underline{\mathbf{z}}) + \log P(\underline{\mathbf{z}}) \right]. \quad (12)$$

Since noise is assumed to be i.i.d Gaussian, from Eqs. (3) and (5) we obtain

$$\begin{aligned} P(\underline{\mathbf{y}}_i \mid \underline{\mathbf{z}}) &= \left[\sum_{i=1}^{p} \log \frac{1}{\left(2\pi\sigma_\eta^2\right)^{\frac{M_1 M_2}{2}}} \times \exp\left\{ -\frac{\| \underline{\mathbf{y}}_i - H_i D\underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} \right\} \right] \\ &= -\sum_{i=1}^{p} \frac{\| \underline{\mathbf{y}}_i - H_i D\underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} - \frac{M_1 M_2}{2} \log\left(2\pi\sigma_\eta^2\right), \end{aligned} \quad (13)$$

where σ_η is the noise variance. Substituting in (12) and using (10) we get,

$$\begin{aligned} \hat{\underline{\mathbf{z}}} &= \arg\max_{\underline{\mathbf{z}}} \left[\sum_{i=1}^{p} -\frac{\| \underline{\mathbf{y}}_i - H_i D\underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} - \sum_{c\in\mathcal{C}} V_c(\underline{\mathbf{z}}) \right] \\ &= \arg\min_{\underline{\mathbf{z}}} \left[\sum_{i=1}^{p} \frac{\| \underline{\mathbf{y}}_i - H_i D\underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} + \sum_{c\in\mathcal{C}} V_c(\underline{\mathbf{z}}) \right] \end{aligned} \quad (14)$$

Substituting Eq. (11) into Eq. (14), the final cost function is obtained as

$$\hat{\underline{\mathbf{z}}} = \arg\min_{\underline{\mathbf{z}}} \left[\sum_{i=1}^{p} \frac{\| \underline{\mathbf{y}}_i - H_i D\underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} + \frac{1}{\lambda} \sum_{k=1}^{N_1} \sum_{l=1}^{N_2} [(z_{k,l} - z_{k,l-1})^2 + (z_{k,l} - z_{k-1,l})^2] \right] \quad (15)$$

The above cost function is convex in terms of the unknown image $\underline{\mathbf{z}}$ and hence a simple gradient descent optimization can be used to minimize it. It may be mentioned here that although the super-resolved image $\underline{\mathbf{z}}$ has been assumed to be an MRF, the low resolution observations $\underline{\mathbf{y}}_i$ do not constitute separate MRFs, and hence a multi-resolution MRF model based super-resolution scheme will not work (see [14] for details).

3.3. Gradient Descent Optimization

It is clear from the cost function of Eq. (15) that it consists of two parts. The first term reflects the error between the observation model and the observed data. If we minimize this term alone, the ill-posedness of the inverse problem could cause excessive noise amplification. The second term is the regularization term which is minimized when $\underline{\mathbf{z}}$ is smooth. The contribution of the two terms are controlled by the noise variance σ_η^2 and the regularization parameter λ. The gradient of (15), at the nth iteration is given by

$$g^{(n)} = \frac{1}{\sigma_\eta^2} \sum_{i=1}^{p} D^T H_i^T \left(H_i D\underline{\mathbf{z}}^{(n)} - \underline{\mathbf{y}}_\mathbf{i} \right) + \frac{G^{(n)}}{\lambda} \quad (16)$$

where $G^{(n)}$ at location (k,l) in the super-resolution lattice is given by

$$G^{(n)}(k,l) = 2\left[4z_{k,l}^{(n)} - z_{k,l-1}^{(n)} - z_{k,l+1}^{(n)} - z_{k-1,l}^{(n)} - z_{k+1,l}^{(n)} \right].$$

The estimate at $(n+1)$th iteration,

$$\underline{\mathbf{z}}^{(n+1)} = \underline{\mathbf{z}}^{(n)} - \alpha g^{(n)}$$

where α is the step size, is computed iteratively until $\|\underline{\mathbf{z}}^{(n+1)} - \underline{\mathbf{z}}^{(n)}\| <$ Threshold. The initial estimate $\underline{\mathbf{z}}^{(0)}$ is chosen as the bilinear interpolation of the available least blurred, low resolution image. It should be noted here that the necessary condition for obtaining the super-resolution image given in Eq. (2) is not applicable here as the super-resolved image $\underline{\mathbf{z}}$ is modeled by an MRF unlike in [7] where $\underline{\mathbf{z}}$ is not represented parametrically. It is the parameteric representation of the super-resolved image $\underline{\mathbf{z}}$ (in terms of the MRF model) that provides the necessary cue for super-resolution.

3.4. Preservation of Discontinuities

Presence or absence of discontinuities conveys important information such as change in surface orientation, depth, texture etc. The concept of line fields on a dual lattice, consisting of sites corresponding to vertical and

horizontal line fields, was introduced in [8]. The horizontal line field $l_{i,j}$ connecting site (i, j) to $(i, j-1)$ aids in detecting a horizontal edge, while the vertical line field $v_{i,j}$ connecting site (i, j) to $(i-1, j)$ helps in detecting a vertical edge. Note that we have chosen $l_{i,j}$ and $v_{i,j}$ to be binary variables in this study. However, one can use continuous variables as well without much changing the problem formulation [15]. The advantage of using continuous variable line fields lies in having a differentiable cost function when a gradient-based optimization method can still be used. The log of the prior distribution in Eq. (10), neglecting the normalizing term, becomes

$$\begin{aligned}\sum_{c\in\mathcal{C}} V_c(\underline{\mathbf{z}}) = \sum_{i,j} \mu[&(z_{i,j} - z_{i,j-1})^2(1 - v_{i,j}) \\ &+ (z_{i,j+1} - z_{i,j})^2(1 - v_{i,j+1}) \\ &+ (z_{i,j} - z_{i-1,j})^2(1 - l_{i,j}) \\ &+ (z_{i+1,j} - z_{i,j})^2(1 - l_{i+1,j})] \\ &+ \gamma[l_{i,j} + l_{i+1,j} + v_{i,j} + v_{i,j+1}] \\ = V(\underline{\mathbf{z}}) \quad &\text{(say)}. \end{aligned} \tag{17}$$

Given a preset threshold, if the gradient at a particular location is above that threshold, the corresponding line field is activated to indicate the presence of a discontinuity. The term multiplying γ provides a penalty for every discontinuity so created. Putting the above expression into Eq. (14), we arrive at the modified cost function

$$\hat{\mathbf{z}} = \arg\min_{\underline{\mathbf{z}}} \left[\sum_{i=1}^{p} \frac{\| \underline{\mathbf{y}}_i - H_i D \underline{\mathbf{z}} \|^2}{2\sigma_\eta^2} + V(\underline{\mathbf{z}}) \right]. \tag{18}$$

When the energy function is non-convex, there is a possibility of the steepest descent type of algorithms getting trapped in a local minima. As shown earlier, our cost function was chosen to be convex. This saved us from the requirement of using a computationally intensive minimization technique like SA. However, on inclusion of line field terms in the cost function to account for discontinuities in the image, the gradient descent technique is liable to get trapped in local minima. We see the similarity between the above cost function and the energy function of the weak membrane formulation [2]. Hence, the GNC algorithm is apt for carrying out the minimization. Although the results indicate an improvement over the gradient descent approach, still better estimates of the super-resolved image $\underline{\mathbf{z}}$ are observed using SA.

4. Experimental Results

Experiments were performed on various images to assert the efficacy of the proposed method. Figure 1 shows two of the five low resolution noisy images of Lena, CT and Pentagon, each of size 64×64, obtained by decimating the respective original images and blurring the decimated images with Gaussian blurs. Although the Gaussian blur has an infinite extent, for purpose of computation we chose the kernel size according to an extent of $\pm 3\sigma$, where σ is the blur parameter. Each low resolution observation contains zero mean Gaussian noise with variance 5.0.

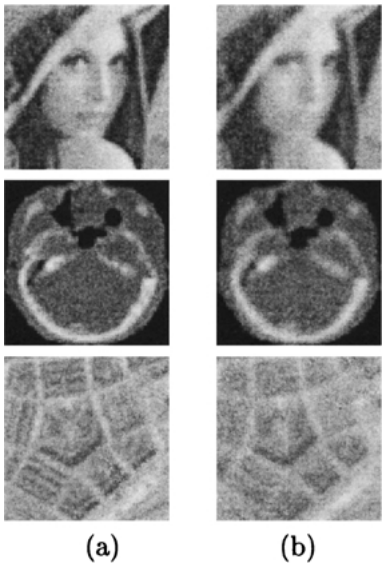

Figure 1. Low resolution, noisy images of Lena, CT and Pentagon with blurs (a) $\sigma = 0.7$ and (b) $\sigma = 1.1$.

Figure 2. Super-resolved Lena image using gradient-descent optimization.

Figure 3. Lena image (a) zero order hold expanded and (b) cubic spline interpolated.

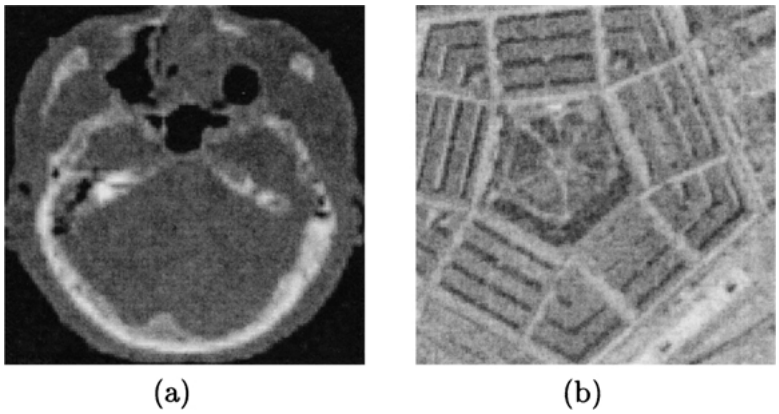

Figure 4. Super-resolved images of (a) CT and (b) Pentagon with 5 low resolution observations, using the gradient-descent method.

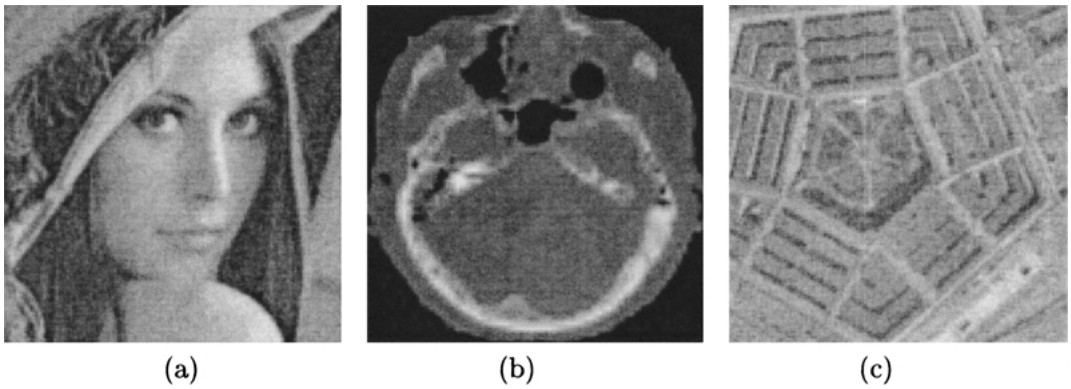

Figure 5. Super-resolved (a) Lena, (b) CT and (c) Pentagon images using only two low resolution observations with $\sigma = 0$ and $\sigma = 0.5$.

First, we present the results of super-resolution using the gradient descent method. As mentioned in Section 3.3, the initial estimate of the high resolution image is the bilinear interpolation of the least blurred observation. The smoothness parameter λ was chosen as 16.75 for the Lena and CT images and 20.0 for the Pentagon image. The step size was initially chosen as 0.1 and was reduced by a factor of 0.99 after each iteration. For large values of λ, the data consistency term in Eq. (15) dominates, producing excessive blockiness in the expanded image. On the other hand, a small value of λ causes over-smoothing. It may be noted that the choice of the value of λ is quite ad-hoc in nature. The value was selected in our study by trial and error method. It may be possible to use a cross-validation technique to arrive at a good choice of λ, but this has not been pursued in this study.

The super-resolved Lena image using the gradient descent optimization scheme is shown in Fig. 2. Results of zero order hold expansion and cubic spline interpolation of the least blurred Lena image are shown in Fig. 3(a) and (b), respectively. The blockiness in the zero-order hold expanded image is clearly discernible, while the spline interpolated image is not only significantly noisy, but, as is expected of splines, it is also blurred due to over-smoothing. On the other hand, the proposed algorithm also does deblurring in addition to

Table 1. Comparison of MSEs of different interpolation schemes.

Method	Lena	CT	Pentagon
ZOH	0.012061	0.899187	0.054042
Cubic spline	0.011870	0.452966	0.052309
Gradient descent	0.003531	0.021216	0.010676
GNC	0.002255	0.021085	0.006962
SA	0.002143	0.007615	0.004186

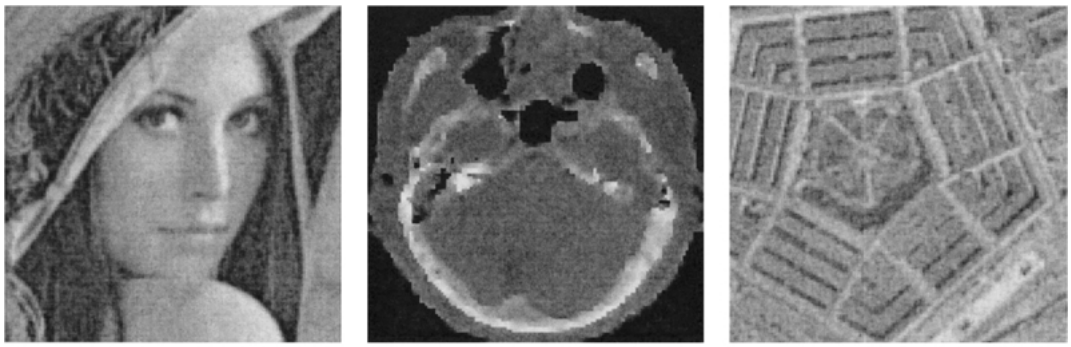

Figure 6. Super-resolved (a) Lena, (b) CT and (c) Pentagon images using the GNC optimization scheme.

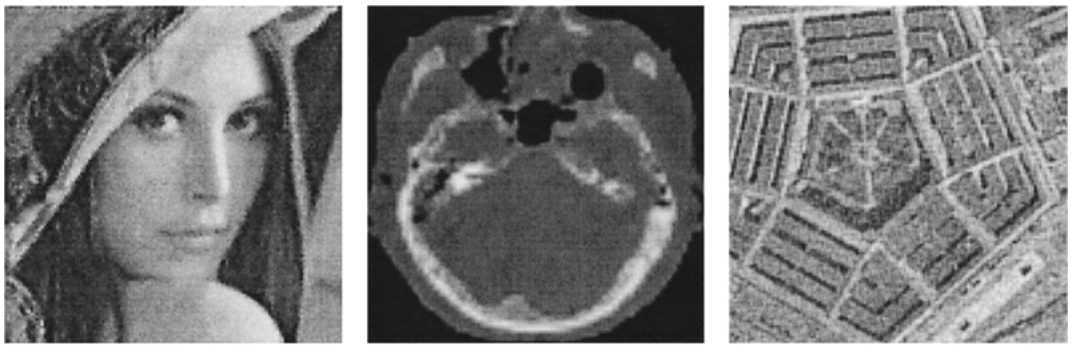

Figure 7. Super-resolved (a) Lena, (b) CT and (c) Pentagon images using simulated annealing (SA).

removing noise and generating a super-resolved image. The super-resolved CT and Pentagon images with 5 low resolution observations using the proposed method are shown in Fig. 4. The gradient descent method was used for the optimization purpose. The cynuses in the bone near the right middle edge of the CT image which are not visible in the low resolution observations shows up clearly in the super-resolved image. The super-resolved Pentagon image contains more details of the circular central part than any of the low-resolution images.

The mean squared error between the original image and generated super resolved image is defined as

$$MSE = \frac{\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(\hat{z}_{i,j} - z_{i,j})^2}{\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}(z_{i,j})^2}. \qquad (19)$$

Table 1 shows the comparison of the MSE for the proposed method with standard methods like zero-order hold and cubic spline interpolation. Notice the significant drop in the MSEs for CT and Pentagon images in going from cubic spline interpolation to the proposed technique using gradient descent optimization technique.

In another experiment, only two low resolution observations, each of Lena, CT and Pentagon were constructed, out of which one was not blurred and the other was blurred with $\sigma = 0.5$. The mean squared errors of the super-resolved Lena, CT and Pentagon images shown in Fig. 5 were 0.003586, 0.021446 and 0.009416 respectively. Compare this to the results given in the third row in Table 1. This is not very different from the results we obtained when all the input images were defocused. Hence, there is no appreciable gain in having focused images in the low resolution ensemble. The proposed technique is, therefore, suitable for low-resolution images that are blurred, since the algorithm inherently performs a deblurring operation. It was observed that there was a marked reduction in the mean square errors till four input images were used; however, for five or more images, the errors did not change significantly, although more number of images helps in smoothing out noise.

Next, we present results of minimization of the modified cost function when line processes are used to preserve discontinuity. As before, we consider 5 low resolution observations. The super-resolved images using GNC as the optimization technique are shown in Fig. 6. The MSE for this method is indicated in Table 1. Visually, there is a significant reduction in noise of the super-resolved image generated using the discontinuity preserving method. In yet another experiment, we carried out the optimization of the modified cost function using SA, but with the output of the GNC algorithm as the initial estimate of the super-resolved image. The super-resolved Lena, CT and Pentagon images from this method are shown in Fig. 7. Notice that the

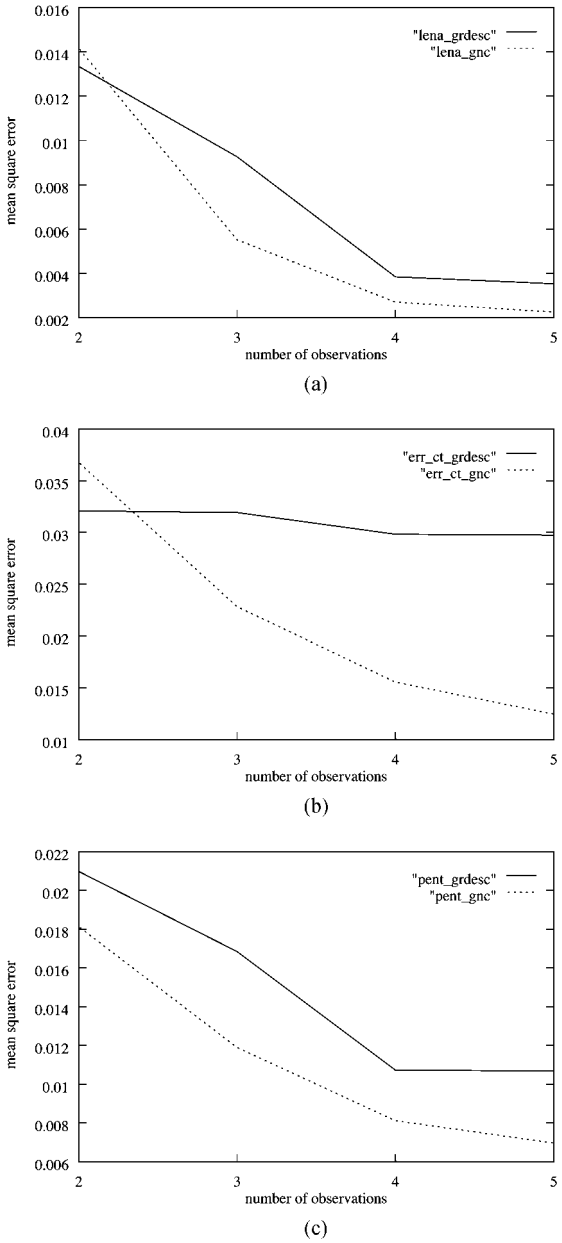

Figure 8. Comparison of mean square errors between the gradient-descent and the discontinuity preserving (GNC) approaches for (a) Lena, (b) CT and (c) Pentagon images, as the number of observations increases.

estimates of the GNC algorithm have undergone further deblurring resulting in a sharper image, e.g. around the eyes of Lena and on the Pentagon image as a whole. We noted earlier that in order to avoid the computational expense of simulated annealing, we opted for a convex cost function by choosing a suitable expression for the clique potentials. However, with incorporation of line fields and optimization using the GNC, which is proven to have a faster convergence than SA, we obtained a better estimate of the super-resolved image. When computational complexity is not an issue, we could go further and use the SA to obtain still better estimates.

Simulations were also carried out to investigate the effect of the number of observations on the quality of the super-resolved image. As shown in Fig. 8, the mean square errors decrease as the number of low resolution observations increases. The plots also illustrate the superiority of the discontinuity preserving method to the gradient descent approach. As noted earlier, the flat nature of the plots for the gradient descent approach implies that the errors do not reduce significantly, although more number of images do contribute to smoothening out noise. On the other hand, increase in the number of images does bring about a substantial reduction in the errors when line fields are included in the cost function.

5. Conclusions

This paper addresses the problem of generating a super-resolution image from a sequence of blurred, decimated and noisy observations of an ideal image. A MAP-MRF approach was used to minimize the function. Comparison with zero order hold and spline interpolation techniques shows that the proposed method is superior. Since there is no relative motion between the observed images, as is the case in most of the previous work in super-resolution, the difficult tasks of image registration and motion estimation are done away with. The proposed technique is fast due to optimization using the gradient descent approach. The errors are seen to level off after about 35 iterations for all the images considered in this paper. Next, the cost function was modified to include line fields to preserve discontinuitites. In addition to significant noise reduction, the sharpness in the image was also observed to be enhanced. This work can also be viewed in the realm of blind image restoration since simultaneous deblurring of the observed images and noise removal is embedded in our method. Future work will involve super-resolved restoration with unknown blurs. This translates to a joint blur identification and super-resolution restoration problem.

References

1. B. Bascle, A. Blake, and A. Zissermann, "Motion deblurring and super-resolution from an image sequence," in *Proc. of European Conf. on Computer Vision*, Cambridge, UK, 1996, Springer-Verlag: Berlin.
2. A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press: Cambridge, MA, 1987.
3. S. Chaudhuri and A.N. Rajagopalan, *Depth from Defocused Images : A Real Aperture Imaging Approach*, Springer-Verlag: New York, 1999.
4. P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "Super-resolved surface reconstruction from multiple images," NASA Ames Research Center, Moffet Field, CA. Technical Report FIA-94-12, 1994.
5. M.-C. Chiang and T.E. Boult, "Local blur estimation and super-resolution," in *Proc. IEEE Conf. Computer Visin and Pattern Recognition*, Puerto Rico, USA, 1997, pp. 821–826.
6. M. Elad and A. Feuer, "Restoration of a single super-resolution image from several blurred, noisy and undersampled measured images," Dept. of Electrical Engg, Technion, Israel Instt. of Technology, Technical Report EE Pub No. 967, 1995.
7. M. Elad and A. Feuer, "Restoration of a single super-resolution image from several blurred, noisy and undersampled measured images," *IEEE Trans. on Image Processing*, Vol. 6, No. 12, pp. 1646–1658, 1997.
8. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of image," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, pp. 721–741, 1984.
9. J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press: New Haven, CT, 1923.
10. M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP:Graphical Models and Image Processing*, Vol. 53, pp. 231–239, 1991.
11. M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion and transparency," *Journal of VCIR*, Vol. 4, pp. 324–335, 1993.
12. S.P. Kim, N.K. Bose, and H.M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Trans. on Accoustics, Speech and Signal Processing*, Vol. 18, No. 6, pp. 1013–1027, 1990.
13. S.P. Kim and W.-Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Trans. on Image Processing*, Vol. 2, pp. 534–539, 1993.
14. S. Krishnamachari and R. Chellappa, "Multiresolution Gauss-Markov random field models for texture segmentation," *IEEE Trans. on Image Processing*, Vol. 6, No. 2, pp. 251–266, 1997.
15. S.Z. Li, *Markov Random Field Modelling in Computer Vision*, Springer-Verlag: Tokyo, 1995.
16. A. Papoulis, "Generalized sampling theorem," *IEEE Trans. on Circuits and Systems*, Vol. 24, pp. 652–654, 1977.
17. A.J. Patti, M.I. Sezan, and A.M. Tekalp, "High resolution image reconstruction from a low resolution image sequence in the

presence of time-varying motion blur," in *Proc. ICIP*, Austin, USA, 1994, pp. 343–347.
18. A.N. Rajagopalan and S. Chaudhuri, "Space-variant approaches to recovery of depth from defocused images," *Computer Vision and Image Understanding*, Vol. 68, No. 3, pp. 309–329, 1997.
19. D. Rajan and S. Chaudhuri, "A generalized interpolation scheme for image scaling and super-resolution," in *Proc. of Erlangen Workshop 99 on Vision, Modelling and Visualization*, University of Erlangen-Nuremberg, Germany, Nov. 1999, pp. 301–308.
20. C.H. Russel, K.J. Barnard, and E.E. Armstrong, "Joint MAP registration and high resolution image estimation using a sequence of undersampled images," *IEEE Trans. on Image Processing*, Vol. 6, No. 12, pp. 1621–1633, 1997.
21. C.H. Russel, K.J. Barnard, J.G. Bognar, E.E. Armstrong, and E.A. Watson, "Joint high resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering*, Vol. 37, No. 1, pp. 247–260, 1998.
22. R.R. Schultz and R.L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Trans. on Image Processing*, Vol. 3, No. 3, pp. 233–242, 1994.
23. R.R. Schultz and R.L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. on Image Processing*, Vol. 5, pp. 996–1011, 1996.
24. H. Shekarforoush, M. Berthod, and J. Zerubia, "3D super-resolution using generalized sampling expansion," in *Proc. Int. Conf. on Image Processing*, Washington D.C., 1995, pp. 300–303.
25. H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman, "Sub-pixel Bayesian estimation of albedo and height," *International Journal of Computer Vision*, Vol. 19, No. 3, pp. 289–300, 1996.
26. C. Srinivas and M.D. Srinath, "A stochastic model based approach for simultaneous restoration of multiple mis-registered images," *SPIE*, Vol. 1360, pp. 1416–1427, 1990.
27. A.M. Tekalp, M.K. Ozkan, and M.I. Sezan, "High resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in *Proc. ICAASP*, San Francisco, USA, 1992, pp. 169–172.
28. B.C. Tom and A.K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration and interpolation of low-resolution images," in *Proc. of Int Conf. Image Processing*, Washington D.C., 1995, pp. 539–542.
29. R.Y. Tsai and T.S. Huang, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processsing*, JAI Press: London, 1984, pp. 317–339.
30. H. Ur and D. Gross, "Improved resolution from sub-pixel shifted pictures," *CVGIP:Graphical Models and Image Processing*, Vol. 54, pp. 181–186, 1992.

Deepu Rajan received the B.E. in Electronics and Communication Engg. from Birla Institute of Technology, Ranchi, India in 1988 and the M.S. in Electrical Engg. from Clemson University, South Carolina, USA in 1990. Currently he is finishing his Ph.D. in the School of Biomedical Engg. at the Indian Institute of Technology-Bombay. Since 1992, he has been a Lecturer in the Department of Electronics, Cochin University of Science and Technology, Cochin, India. His research interests are in image and video processing, multimedia communication and neural networks.

Subhasis Chaudhuri was born in Bahutali, India. He received his B. Tech. degree in Electronics and Electrical Communication Engineering from the Indian Institute of Technology, Kharagpur in 1985. He received the M.S and Ph.D. degrees, both in Electrical Engineering, respectively, from the University of Calgary, Canada and the University of California, San Diego. He joined the IIT, Bombay in 1990 as an assistant professor and is currently serving as a professor. He has also served as a visiting professor at the University of Erlangen-Nuremberg, Germany during the summer of 1996. He is a fellow of the Alexander von Humboldt Foundation, Germany and the IETE, India. He is the co-author of the book 'Depth from Defocus: A Real Aperture Imaging Approach', published by Springer, NY. He is currently editing a book on super-resolution imaging to be published by Kluwer Academic. His research interests include image processing, computer vision, multimedia.