Robust and computationally efficient superresolution algorithm

Kaggere V. Suresh and Ambasamudram N. Rajagopalan

Department of Electrical Engineering, Image Processing and Computer Vision Laboratory, Indian Institute of Technology Madras, Chennai 600 036, India

Received May 5, 2006; revised August 14, 2006; accepted October 9, 2006; posted October 20, 2006 (Doc. ID 70650); published March 14, 2007

Superresolution is the process of combining information from multiple subpixel-shifted low-resolution images to form a high-resolution image. It works quite well under ideal conditions but deteriorates rapidly with inaccuracies in motion estimates. We model the original high-resolution image as a Markov random field (MRF) with a discontinuity adaptive regularizer. Given the low-resolution observations, an estimate of the superresolved image is obtained by using the iterated conditional modes (ICM) algorithm, which maximizes the local posterior conditional probability sequentially. The proposed method not only preserves edges but also lends robustness to errors in the estimates of motion and blur parameters. We derive theoretically the neighborhood structure for the posterior distribution in the presence of warping, blurring, and downsampling operations and use this to effectively reduce the overall computations. Results are given on synthetic as well as real data to validate our method.

OCIS codes: 100.6640, 100.3010.

1. INTRODUCTION

Superresolution is a signal processing technique to obtain a high-resolution (HR) image from multiple low-resolution (LR) observations. The effect of LR sensors is to cause aliasing of high-frequency components and local blurring. Motion-based superresolution algorithms use information from subpixel-shifted LR observations to reconstruct a HR image.[1–4] It is also possible to perform motion-free superresolution in which the idea is to generate a HR image from a set of defocused and downsampled observations blurred to different extents.[5–8] Yet another group of superresolution algorithms called learning-based methods use specific information about the class of images to be superresolved.[9,10]

In this paper, we focus on motion-based superresolution for which many algorithms exist. These can be broadly divided into two classes: frequency-domain methods and spatial-domain methods. Huang and Tsai[11] proposed a frequency-domain formulation based on the shift and aliasing properties of the continuous and discrete Fourier transforms. The observed images are modeled as undersampled images of a static unknown scene. Kim *et al.*[12] extended the formulation in Ref. 11 to consider observation noise as well as effects of spatial blurring. Ur and Gross[13] proposed a method based on the generalized sampling theorem. Though frequency-domain-based methods are simple and computationally attractive, they suffer from the drawback of being able to deal only with global translational motion.

Superresolution methods based on spatial-domain formulation[14] can accommodate general motion models and are also suitable for the inclusion of *a priori* constraints on the image. A large class of spatial-domain superresolution methods use a simple and powerful simulate-and-correct approach commonly known as iterated backprojection (IBP).[15] These methods assume that some estimate of the HR image is known. Given the knowledge of the imaging process relating the scene to the observations, the output of the imaging system is simulated from an initial estimate of the HR image. The simulated images are then compared with the observed data, and the error is backprojected to update the estimate of the HR image.

Superresolution image reconstruction is basically an ill-posed inverse problem. Minimization of the error between the simulated images and the observations does not necessarily imply a good solution. Methods that utilize *a priori* constraints to obtain regularized solutions are required. Bayesian methods are well suited, as they provide a natural framework for inclusion of *a priori* constraints in the form of a prior probability function on the unknown. Stevenson and Schultz[16] used a Huber–Markov random-field (HMRF) prior model, which imposes smoothness constraints below a fixed threshold value while allowing the restoration of visually important edge features. Hardie *et al.*[17] presented a maximum *a posteriori* (MAP) superresolution procedure using a Gaussian MRF (GMRF) prior. Elad and Feuer[5] addressed the superresolution problem from the point of view of maximum likelihood (ML), MAP, and projection onto convex sets (POCS).

Most motion-based superresolution algorithms assume that the motion and blur estimates are available or use some method to estimate them. When these parameters are assumed or estimated incorrectly, the quality of the superresolved image suffers. Accurate registration among the LR images is very important for good reconstruction. However, subpixel-level motion estimation is a difficult problem. Zomet *et al.*[18] proposed a median-based estimator to discard measurements that are inconsistent with

the imaging model. Lee and Kang[19] proposed a regularized adaptive algorithm in which the registration error noise is modeled as Gaussian. Farsiu *et al.*[20] proposed an alternate approach using L_1 norm minimization and a regularization prior. In Refs. 16 and 17 noise due to motion error is incorporated as *a priori* information within the smoothness prior and MAP estimation is used to obtain the solution.

In this paper, we present a computationally efficient method for superresolution using a discontinuity adaptive MRF (DAMRF) model. Our motivation behind using an MRF model for the original HR image is to provide robustness to errors in motion and blur estimates. A discontinuity adaptive (DA) regularizer is proposed in which the degree of interaction between pixels across edges is adjusted adaptively to preserve discontinuities. In a genuine MRF algorithm, no two neighboring sites should be updated simultaneously.[21] We derive the MAP estimate of the original image using the iterated conditional modes (ICM)[22] algorithm, which maximizes the local conditional probabilities sequentially. However, such an update procedure is computationally very intensive. To address this issue, we prove an important theorem that, given the observations, the posterior distribution is Markov for the motion superresolution problem and formally derive the exact posterior neighborhood structure. The novelty of this theorem lies in the fact that it enables us to propose a computationally efficient ICM algorithm for superresolution. The performance of the proposed method is found to be superior to existing methods in not only preserving edges but also lending robustness in the presence of errors in the estimates of motion and blur parameters.

The paper is organized as follows: In section 2, we discuss MAP formulation of the superresolution problem. Section 3 describes a DAMRF prior. In Section 4, we propose minimization using the ICM algorithm. We derive the exact posterior neighborhood structure and show how locality can be used for computational speedup. Experimental results are given in Section 5, while section 6 concludes the paper.

2. MAP FORMULATION

The sources of degradations during the process of acquisition of a sequence of LR images can be divided into three parts: geometric transformation, blur, and downsampling. The geometric transformation of the original image is usually assumed to be simple translational motion. The blurring can be due to camera (defocus blur) and relative motion between object and camera (motion blur). The relationship between a lexicographically ordered LR observation and the original HR image is given by

$$\underline{y}_r = DH_r^{\text{cam}} W_r H_r^{\text{mot}} \underline{x} + \underline{n}_r, \qquad 1 \leq r \leq m, \tag{1}$$

where

- $\underline{x}$: Original HR image ($N_1N_2 \times 1$).
- $\underline{y}_r$: rth LR observation ($M_1M_2 \times 1$).
- D: Downsampling matrix ($M_1M_2 \times N_1N_2$).
- H_r^{cam}: Camera defocus blur matrix for rth frame ($N_1N_2 \times N_1N_2$).
- W_r: Geometric warping matrix for rth frame ($N_1N_2 \times N_1N_2$).
- H_r^{mot}: Motion blur matrix for rth frame ($N_1N_2 \times N_1N_2$).
- $\underline{n}_r$: Noise in the rth frame.
- $\overline{m}$: Number of LR observations.

Equation (1) can be expressed in matrix-vector form as

$$\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_m \end{pmatrix} = \begin{pmatrix} DH_1^{\text{cam}} W_1 H_1^{\text{mot}} \\ DH_2^{\text{cam}} W_2 H_2^{\text{mot}} \\ \vdots \\ DH_m^{\text{cam}} W_m H_m^{\text{mot}} \end{pmatrix} \underline{x} + \begin{pmatrix} \underline{n}_1 \\ \underline{n}_2 \\ \vdots \\ \underline{n}_m \end{pmatrix}. \tag{2}$$

The MAP estimate of the superresolved image $\underline{x}$ given m LR images is

$$\hat{\underline{x}} = \arg\max_{\underline{x}} \{P(\underline{x} | \underline{y}_1, \ldots, \underline{y}_m)\}. \tag{3}$$

Using Baye's rule, this can be rewritten as

$$\hat{\underline{x}} = \arg\max_{\underline{x}} \left\{ \frac{P(\underline{y}_1, \ldots, \underline{y}_m | \underline{x}) P(\underline{x})}{P(\underline{y}_1, \ldots, \underline{y}_m)} \right\}. \tag{4}$$

Since the denominator is not a function of $\underline{x}$,

$$\hat{\underline{x}} = \arg\max_{\underline{x}} \{P(\underline{y}_1, \ldots, \underline{y}_m | \underline{x}) P(\underline{x})\}. \tag{5}$$

Taking the logarithm of the posterior probability, the MAP estimate of $\underline{x}$ is given by

$$\hat{\underline{x}} = \arg\max_{\underline{x}} \{\log[P(\underline{y}_1, \ldots, \underline{y}_m | \underline{x})] + \log P(\underline{x})\}. \tag{6}$$

The above MAP formulation allows us to incorporate prior knowledge about $\underline{x}$ for improving robustness during reconstruction.

3. DAMRF MODEL

From Eq. (6), we note that we must specify a prior probability density for the original HR image X. We model the superresolved image to be estimated as a MRF because it provides a foundation for the characterization of contextual constraints and the densities of the probability distributions of interacting features in images.

Let F be a random field over an $N \times N$ lattice of sites $L = (i,j): 1 \leq i,j \leq N$. The random field F is said to be an MRF on L with respect to a neighborhood system η if

1. $P(F=f) > 0, \ \forall f \in \mathcal{F}$,
2. $P[F_{i,j}=f_{i,j} | F_{k,l}=f_{k,l}, \forall (k,l) \neq (i,j)] = P[F_{i,j}=f_{i,j} | F_{k,l}=f_{k,l}, (k,l) \in \eta_{i,j}]$,

where $\eta_{i,j}$ is the neighborhood of the site (i,j) and $\mathcal{F}$ denotes the configuration space. It is natural to expect that the image intensity at a pixel will not depend on the image data outside its neighborhood when the image data on its neighborhood are given. MRF image models even with first-order neighborhood system are known to be powerful.[23]

The practical use of MRF models can be largely ascribed to the equivalence between MRFs and Gibbs random field (GRF) established by Hammersely and

Clifford.[24] The theorem states that F is an MRF on L with respect to neighborhood η if and only if F is a Gibbs random field on L. That is,

$$P[\underline{F}=\underline{f}]=(1/Z)\exp\{-U(\underline{f})\}, \tag{7}$$

where Z is the partition function given by $Z=\Sigma_f\exp\{-U(\underline{f})\}$ and $U(\underline{f})$ is the energy function, which is given by

$$U(\underline{f})=\sum_{c\in C}V_c(\underline{f}). \tag{8}$$

Here, c is called the clique of the pair (L,η), which is a subset of sites in L in which all pairs of sites are mutual neighbors. The set C is the set of all cliques. Since we model the HR image X as an MRF, we can write

$$P[\underline{X}=\underline{x}]=(1/Z)\exp\{-U(\underline{x})\}, \tag{9}$$

where

$$U(\underline{x})=\sum_{c\in C}V_c(\underline{x}). \tag{10}$$

The choice of the clique potential $V_c(\underline{x})$ is crucial, as it embeds important prior information about the image to be reconstructed. The prior model can be chosen as

$$\sum_{c\in C}V_c(\underline{x})=\sum_{c\in C}g(d_c\underline{x}),$$

where $d_c\underline{x}$ is a local spatial activity measure of the image and has a small value in smooth regions and a large value at edges.

A common choice for the prior is a Gauss–Markov random-field model,[17] which has the form

$$g(n)=n^2.$$

This image model can result in a blurred estimate of X, particularly along edges, owing to oversmoothing. Geman and Geman[24] introduced line fields for preserving edges. But the use of line fields which are binary valued makes the problem nondifferentiable. Stevenson and Schultz[16] used a discontinuity preserving model of the form

$$g(n)=\begin{cases}n^2 & |n|\leqslant T\\ 2T(|n|-T)+T^2 & |n|>T\end{cases},$$

where T is the threshold parameter separating the quadratic and linear regions. This choice of $g(n)$ leads to the HMRF. The threshold, which is dependent on factors like image content and inherent noise, has to be tuned for each individual case, which is impractical. If fixed at low values, the threshold lets in noise, whereas at high values it penalizes weak edges.

We propose to use a DAMRF model in which the degree of interaction between pixels across edges is adjusted adaptively in order to preserve discontinuities. A necessary condition for any regularization model to be adaptive to discontinuities[21] is

$$\lim_{n\to\infty}|g'(n)|=\lim_{n\to\infty}|2nh(n)|=C, \tag{11}$$

where n is the difference between neighboring pixel values and $C\in[0,\infty)$ is a constant. We propose to choose $g(n)$ as

$$g(n)=\gamma-\gamma e^{-n^2/\gamma}. \tag{12}$$

Figure 1 shows the function defined by Eq. (12). It is convex in the band $B_\gamma=(-\sqrt{\gamma/2},\sqrt{\gamma/2})$ and nonconvex outside. The DA function allows the smoothing strength to increase monotonically as n increases within the band B_γ, thus smoothing out noise. Outside this band, smoothing decreases as n increases, thereby preserving the discontinuities. Using this DA function and assuming a first-order MRF neighborhood, we have

$$\begin{aligned}\sum_{c\in C}V_c(\underline{x})=\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}&4*\gamma-\gamma\exp[-(x(i,j)-x(i,j-1))^2/\gamma]\\&-\gamma\exp[-(x(i,j)-x(i,j+1))^2/\gamma]-\gamma\exp[-(x(i,j)\\&-x(i-1,j))^2/\gamma]-\gamma\exp[-(x(i,j)-x(i+1,j))^2/\gamma],\end{aligned} \tag{13}$$

$$P(\underline{X}=\underline{x})=(1/Z)\exp\left\{-\sum_{c\in C}V_c(\underline{x})\right\}, \tag{14}$$

where $\Sigma_{c\in C}V_c(\underline{x})$ is given by Eq. (13). Using the observation model in Eq. (1) and the fact that the noise fields are statistically independent of X and as well as each other, we have

$$\begin{aligned}&P[\underline{Y_1}=\underline{y_1},\ldots,\underline{Y_m}=\underline{y_m}|\underline{X}=\underline{x}]\\&\quad=\frac{1}{(2\pi\sigma^2)^{m(M_1M_2/2)}}\exp\left\{-\sum_{r=1}^{m}\frac{\|\underline{y_r}-DH_r^{\mathrm{cam}}W_rH_r^{\mathrm{mot}}\underline{x}\|^2}{2\sigma^2}\right\},\end{aligned} \tag{15}$$

where σ^2 is the variance of the observation noise. From Eqs. (14) and (15), the posterior distribution can be written as

$$P[\underline{X}=\underline{x}|\underline{Y_1}=\underline{y_1},\ldots,\underline{Y_m}=\underline{y_m}]=K\exp\{-U^p(\underline{x})\}, \tag{16}$$

where the posterior energy function $U^p(\underline{x})$ is given by

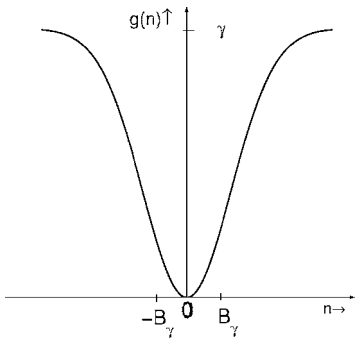

Fig. 1. Discontinuity adaptive function.

$$U^p(\underline{x}) = \sum_{r=1}^{m} \frac{\|\underline{y}_r - DH_r^{\text{cam}} W_r H_r^{\text{mot}} \underline{x}\|^2}{2\sigma^2} + \lambda \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 4 * \gamma - \gamma \exp[-(x(i,j) - x(i,j-1))^2/\gamma] - \gamma \exp[-(x(i,j) - x(i,j+1))^2/\gamma] - \gamma \exp[-(x(i,j) - x(i-1,j))^2/\gamma] - \gamma \exp[-(x(i,j) - x(i+1,j))^2/\gamma], \tag{17}$$

and K is a normalizing constant. The parameter λ corresponds to the weight of the smoothness term. Thus, computing the MAP estimate of X is equivalent to minimizing $U^p(\underline{x})$.

4. ICM ALGORITHM

In this section, we consider the problem of minimizing the energy function $U^p(\underline{x})$ sequentially. It is a well-known fact that in a genuine MRF algorithm, no two neighboring sites should be updated simultaneously.[21] The ICM is a deterministic algorithm that maximizes local conditional probabilities sequentially using a "greedy" strategy. Given the LR observations y_r and the labels $x_{\mathcal{F}-\{i\}}$ of the HR image X, we propose to use ICM to sequentially update x_i^k to x_i^{k+1} by maximizing $P(x_i|y_r, x_{\mathcal{F}-\{i\}})$, the posterior probability. The initial estimate of the superresolved image can be taken as the average of the bilinearly upsampled and aligned input images. Each pixel x_i^k is updated to a new value such that the energy $U^p(\underline{x})$ is minimum over all values of x_i. Because the DA function (Fig. 1) is nonconvex, the overall cost function is also nonconvex. To alleviate the effect of local minima, we use deterministic annealing for optimization. The overall cost function is made approximately convex initially by choosing a large value of γ. This is decreased gradually after every iteration until convergence. As the iterations progress, the effect of smoothness decreases, thereby preserving sharp features in the image. The algorithm is given below.

Calculate the initial estimate $X(0)$ as the average of the bilinearly upsampled and aligned images.

for $l=1$ to k, do
begin
  for $i=1:N_1$, $j=1:N_2$, do
  begin
    old_value$=X_{i,j}$, new_value$=0$;
    REPEAT
      $X_{i,j}$=new_value;
    if $U^p(\underline{x}(\text{new\_value})) \leq U^p(\underline{x}(\text{old\_value}))$
      then set $\underline{x}$(old_value)=$\underline{x}$(new_value);
        $U^p(\underline{x}(\text{old\_value}) = U^p(\underline{x}(\text{new\_value})))$;
    else $X_{i,j}$=old_value;
    new_value=new_value$+\delta$;
    UNTIL new_value$>$maximum allowed gray level $=255$
  end
$\gamma_{\text{new}} = r * \gamma_{\text{old}}$;
end
Set $\hat{X}=X$(new);

δ is the increment in gray value, r is a constant less than unity, and k is the number of iterations.

It must be noted that a direct implementation of the ICM algorithm will be computationally very intensive, as we need to calculate the energy $U^p(\underline{x})$ over the whole image for every new value of a pixel. To arrive at a computationally efficient algorithm, it is important to examine whether the posterior distribution possesses the locality property. The posterior distribution must have a reasonable neighborhood structure as an MRF so that it can accommodate the computational load of the problem.

Before we determine the neighborhood for the posterior distribution, we note that warping, blurring, and downsampling operations depend on a specific neighborhood. Usually, the defocus blur is modeled as Gaussian with a finite support ($\pm 3\sigma_b$), where σ_b is the blur parameter. The motion blur also has a finite support given by its length L. Warping has a finite support of 2 (each pixel is a weighted average of its diagonal neighbors), while downsampling has a support of q, where q is the resolution factor.

We next derive the neighborhood for the posterior distribution mathematically.

Theorem: The posterior neighborhood corresponding to the site (i,j) in the superresolved image is given by

$$\eta_{i,j}^p = \eta_{i,j}^x \bigcup_{r=1}^{n} \left\{ \bigcup_{(k,l) \in \{A - A_r\}} \zeta_{k,l}^{y_r} \right\},$$

where $\eta_{i,j}^x$ is the neighborhood in the superresolved image corresponding to the original MRF image model, $\zeta_{k,l}^{y_r}$ is the set of pixels in the superresolved image that affects the LR image y_r at site (k,l), $A=\{(k,l): 1 \leq (k,l) \leq M_1 M_2\}$, and $A_r = \{(k,l): (i,j) \notin \zeta_{k,l}^{y_r}\}$.

Proof: The conditional probability of $X_{i,j}$ given all the m observations is

$$\begin{aligned} P[X_{i,j} = x_{i,j}; 1 \leq (i,j) \leq N_1 N_2 | \underline{Y}_1 = \underline{y}_1, \ldots, \underline{Y}_m = \underline{y}_m] \\ = P[X_{i,j} = x_{i,j} | X_{k,l} = x_{k,l}; 1 \leq (k,l) \leq N_1 N_2, (k,l) \\ \neq (i,j); \underline{Y}_1 = \underline{y}_1, \ldots, \underline{Y}_m = \underline{y}_m] P[X_{k,l} \\ = x_{k,l}; 1 \leq (k,l) \leq N_1 N_2, (k,l) \\ \neq (i,j) | \underline{Y}_1 = \underline{y}_1, \ldots, \underline{Y}_m = \underline{y}_m]. \end{aligned}$$

Using Eq. (16), we obtain

$$\begin{aligned} P[X_{i,j} = x_{i,j} | X_{k,l} = x_{k,l}; 1 \leq (k,l) \leq N_1 N_2, (k,l) \\ \neq (i,j); \underline{Y}_1 = \underline{y}_1, \ldots, \underline{Y}_m = \underline{y}_m] \\ = \frac{\exp\{-U^p(\underline{x})\}}{\sum_{x_{i,j} = \text{all possible levels}} \exp\{-U^p(\underline{x})\}}. \end{aligned} \tag{18}$$

Define the vectors

$$\underline{\Psi}^r = \frac{1}{\sqrt{2}\sigma}(\underline{y}_r - DH_r^{\text{cam}} W_r H_r^{\text{mot}} \underline{x}), \qquad 1 \leq r \leq m. \tag{19}$$

By notation $\psi_{i,j}$ is the $[(i-1)M_1 M_2 + j]$th element of vector $\underline{\psi}$. The posterior energy function can then be written as

$$U^p(\underline{x}) = \sum_{r=1}^{m} \sum_{1 \le (k,l) \le M_1 M_2} \overline{\psi^r_{k,l}}^2 + \sum_{c \in C} V_c(\underline{x}). \quad (20)$$

We can then decompose the energy function as

$$U^p(\underline{x}) = \sum_{r=1}^{m} \sum_{\{A-A_r\}} \overline{\psi^r_{k,l}}^2 + \sum_{c \in C,(i,j) \in c} V_c(\underline{x}) + \sum_{r=1}^{m} \sum_{\{A_r\}} \overline{\psi^r_{k,l}}^2 + \sum_{c \in C,(i,j) \notin c} V_c(\underline{x}). \quad (21)$$

Substituting Eq. (21) into Eq. (18) and canceling terms common to the numerator and the denominator, we obtain

$$P[X_{i,j} = x_{i,j} | X_{k,l} = x_{k,l}; 1 \le (k,l) \le N_1 N_2, (k,l) \ne (i,j); \overline{Y_1} = \overline{y_1}, \ldots, \overline{Y_m} = \overline{y_m}] = \frac{\exp\left\{ -\sum_{r=1}^{m} \sum_{\{A-A_r\}} \overline{\psi^r_{k,l}}^2 - \sum_{c \in C,(i,j) \in c} V_c(\underline{x}) \right\}}{\sum_{x_{i,j}} \exp\left\{ -\sum_{r=1}^{m} \sum_{\{A-A_r\}} \overline{\psi^r_{k,l}}^2 - \sum_{c \in C,(i,j) \in c} V_x(\underline{x}) \right\}},$$

where $\{A-A_r\} = \{(k,l)$ in the LR observation $y_r : (i,j) \in \zeta^{y_r}_{k,l}\}$. Each of these (k,l) in y_r has a finite neighborhood $\zeta^{y_r}_{k,l}$ in the HR image. Hence, the posterior neighborhood structure corresponding to site (i,j) is

$$\eta^p_{i,j} = \eta^x_{i,j} \cup_{r=1}^{m} \left\{ \cup_{(k,l) \in \{A-A_r\}} \zeta^{y_r}_{k,l} \right\}.$$

For purpose of illustration, we show an example corresponding to the following case: HR image of dimension 8×8 pixels, $q=2$, $\sigma_b = 1/3$ (a 3×3 kernel), and $L=3$. We assume subpixel warping in the forward direction by an amount (δ_x, δ_y). Since the generation of LR observations from HR image involves blurring due to motion, warping, blurring due to camera defocus, and downsampling, we compute the neighborhood $\zeta^{y_r}_{k,l}$ by examining the support for each operation. The shaded portion in Fig. 2(a) shows the site (k,l) in the LR image y_r, and Figs. 2(b)–2(e) show the support for each of the operations. Note that $\zeta^{\text{samp}}_{k,l}$, $\zeta^{\text{cam}}_{k,l}$, $\zeta^{\text{warp}}_{k,l}$, and $\zeta^{\text{mot}}_{k,l}$ are the supports for downsampling, camera defocus blur, warping, and blurring due to motion, respectively. Figure 2(e) is the overall neighborhood $\zeta^{y_r}_{k,l}$ for the site (k,l).

In Fig. 3, for site (i,j) we indicate the sequence of steps leading to the set $\{A-A_r\}$ in the LR observation y_r. From Fig. 2(e) we note that each (k,l) has a finite neighborhood $\zeta^{y_r}_{k,l}$ in the HR image. The union of all these $\zeta^{y_r}_{k,l}$'s over all the LR observations along with the original MRF neighborhood yields the overall posterior neighborhood. This neighborhood is finite and is much smaller than the actual image dimensions.

Based on the posterior neighborhood, the ICM algorithm proposed earlier in this section can be implemented very efficiently. Note that the blurring operation due to motion involves L multiplications and $(L-1)$ additions; warping using bilinear interpolation needs 8 multiplications and 7 additions; blurring due to defocusing needs $(\Delta+1)^2$ multiplications and $((\Delta+1)^2-1)$ additions, where

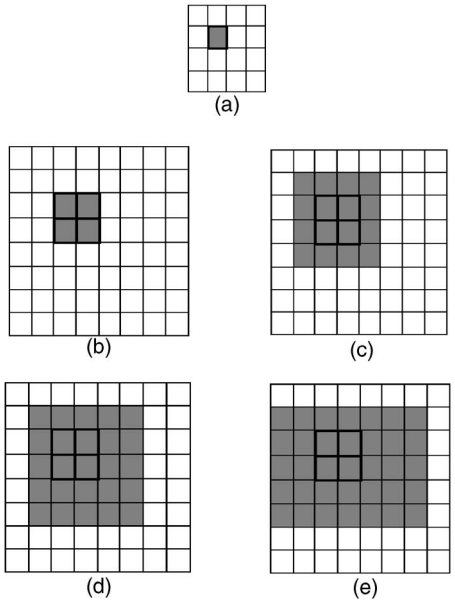

Fig. 2. Pictorial representation of the neighborhood for different operations. (a) Site (k,l) in LR image of size 4×4 pixels. Neighborhood structure after (b) downsampling ($\zeta^{\text{samp}}_{k,l}$), (c) defocus blurring ($\zeta^{\text{cam}}_{k,l}$), (d) warping ($\zeta^{\text{warp}}_{k,l}$), and (e) motion blurring ($\zeta^{\text{mot}}_{k,l}$).

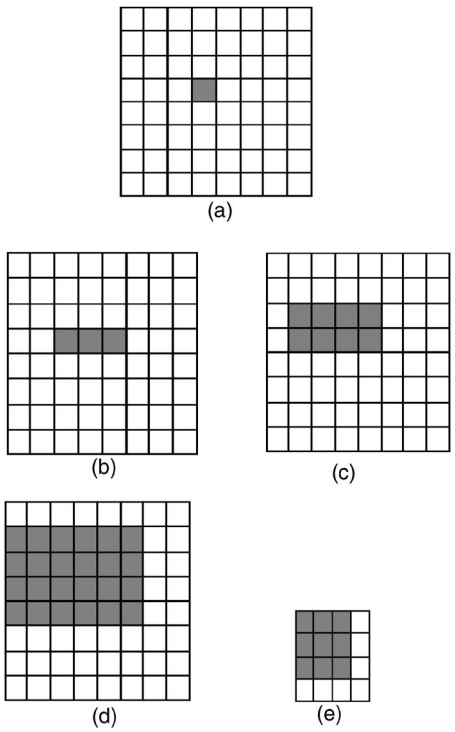

Fig. 3. Illustration of the set $\{A-A_r\}$ for the case of $q=2$, $\sigma_b = 0.33$, and $L=3$. (a) Pixel (i,j) in HR image of size 8×8 pixels. Neighborhood structure after (b) motion blurring, (c) warping, (d) defocus blurring, and (e) downsampling.

$\Delta = 2\,\text{ceil}(3\sigma_b)$; and downsampling involves 1 multiplication and (q^2-1) additions. When the pixel $x(i,j)$ is changed, the number of computations required to determine the first term in $U^p(\underline{x})$ [Eq. (17)] without the locality criterion is given in column 2 of Table 1. [Note that computation of the prior term in $U^p(\underline{x})$ effectively involves

only the four neighbors of $x(i,j)$]. For each of the operations in Table 1, the first and the second row correspond to the required number of multiplications and additions, respectively. For large N_1 and N_2, the computational load for ICM (without locality) will be intensive. However, as shown in Fig. 3, the pixels that are actually affected are, in fact, very few. The computational requirement for the proposed ICM algorithm that utilizes this locality constraint is significantly less (column 3 of Table 1), rendering it computationally attractive.

5. EXPERIMENTAL RESULTS

In this section, we demonstrate the effectiveness of the proposed DAMRF method for the superresolution problem. In all the experiments, we consider resolution improvement by a factor of 2. The method is tested on synthetic as well as real images. The values chosen for the various parameters in the ICM algorithm were $\lambda=0.001$, $\gamma=300$, $r=0.95$, $k=5$, and $\delta=10$. Initially, we assumed that accurate motion and blur estimates are available. In the later part, we analyzed the effect of errors in the estimates of these parameters.

We first compare the performance of the proposed algorithm (which uses a model for the HR image) with those that do not use a model. We start with a single high-quality image of size 140×140 pixels shown in Fig. 4(a), from which we generate four geometrically warped, blurred, and downsampled images of size 70×70 pixels. To each LR image, independent white Gaussian noise of variance 5 was added. Warping to the subpixel level was done using bilinear interpolation. We used a Gaussian kernel with standard deviation 0.33 for camera defocus blur and a horizontal 1-D window of length 3 for motion blur. Figure 4(b) shows one of the degraded LR images, and Fig. 4(c) shows the bilinearly upsampled image. Figures 4(d)–4(f) show the reconstructed images using least-squares (LS) technique,[25] regularized least squares (RLS),[25] and the ICM algorithm based on DAMRF, respectively. Note that the proposed method yields the best re-

Table 1. Comparison of Computational Requirements

Operation	ICM (Without Locality)	ICM (With Locality)
Motion blurring	$N_1N_2\cdot L$	$L\cdot L$
	$N_2N_2\cdot(L-1)$	$L\cdot(L-1)$
Warping	$N_1N_2\cdot 8$	$2(L+1)\cdot 8$
	$N_1N_2\cdot 7$	$2(L+1)\cdot 7$
Defocus blurring	$N_1N_2\cdot(\Delta+1)^2$	$(2+\Delta)(L+1+\Delta)\cdot(\Delta+1)^2$
	$N_1N_2\cdot[(\Delta+1)^2-1]$	$(2+\Delta)(L+1+\Delta)\cdot[(\Delta+1)^2-1]$
Downsampling	$\frac{N_1N_2}{q^2}\cdot 1$	$\left[\text{ceil}\left(\frac{2+\Delta}{q}\right)+1\right]\left[\text{ceil}\left(\frac{L+1+\Delta}{q}\right)+1\right]\cdot 1$
	$\frac{N_1N_2}{q^2}\cdot(q^2-1)$	$\left[\text{ceil}\left(\frac{2+\Delta}{q}\right)+1\right]\left[\text{ceil}\left(\frac{L+1+\Delta}{q}\right)+1\right]\cdot(q^2-1)$

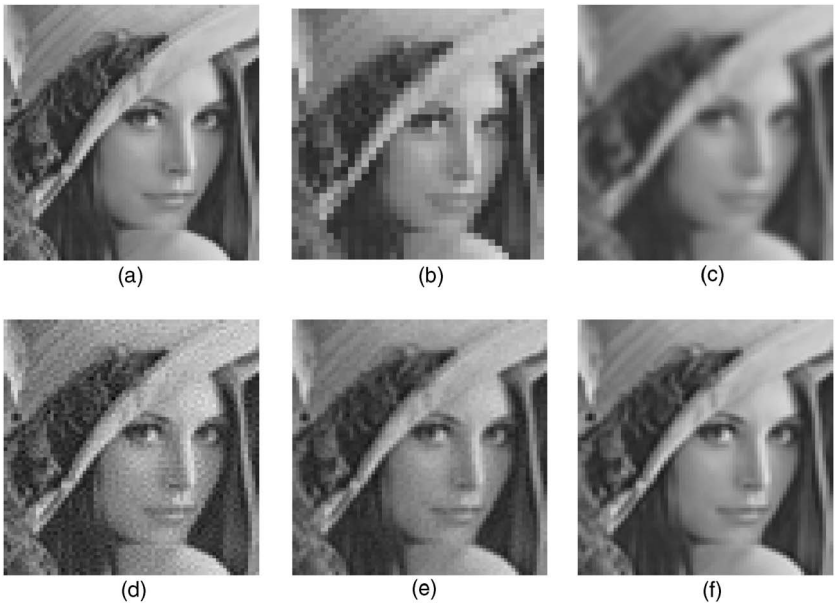

Fig. 4. Comparison among LS, RLS, and DAMRF. (a) Original, (b) LR observation, and (c) bilinear interpolation result. Superresolved image corresponding to (d) LS, (e) RLS, and (f) DAMRF.

Table 2. Comparison among LS, RLS, and DAMRF Methods

Method	$\sigma^2=2.25$	$\sigma^2=5$	$\sigma^2=16$
Least squares	110.47	261.21	927.8
Regularized least squares	50.07	56.74	81.50
DAMRF	19.37	26.45	46.03

sult [Fig. 4(f)]. The mean-squared error (MSE) per pixel is given in Table 2 for the three methods for different noise variances. The MSE is lowest for the proposed algorithm.

Next, we compare the performance of the proposed technique that uses DAMRF with that of GMRF[17] and HMRF.[16] We start with a HR image of size 100×100 pixels as shown in Fig. 5(a). Four LR observations were obtained synthetically as explained in the previous experiment. We assumed a defocus blur with $\sigma_b = 0.33$ and noise variance 16. Figure 5(b) shows one of the LR observations, and Fig. 5(c) is the result obtained using bilinear interpolation. Figures 5(d)–5(f) show reconstructed images using GMRF, HMRF, and the proposed DAMRF method, respectively. A fine-tuned threshold value of $T = 1$ was used for HMRF. The GMRF method reduces noise but leads to oversmoothing. In comparison with GMRF, HMRF yields slightly better results. However, the reconstructed image output using DAMRF is the best. It is less noisy and has more distinctly defined edges. This is clearly evident from the error images shown in Figs. 5(g)–5(i). Among the three models, the error between the original and the reconstructed image using the proposed DAMRF model is the lowest.

In the next experiment, we evaluate the performance of our method in the presence of parameter estimation errors. First, we performed an experiment to evaluate the performance of the algorithm when the motion estimates are in error. Although the subpixel motion information of the simulated LR images is known accurately, we assumed that the estimation of the subpixel motion is incorrect. Four LR observations were generated from a single HR "pepper" image with motion parameters (0,0), (0,0.5), (0.5,0), and (0.5,0.5). During reconstruction, the above motion parameters were deliberately fed incorrectly as (0,0), (0,0.4), (0.4,0), and (0.5,0.4). Figure 6(a) shows the original image, and Fig. 6(b) shows the output of the algorithm. Note that despite motion errors, the superresolved image looks good. Next, we considered error in blur estimation. While generating the LR images, we blurred the images with a Gaussian kernel of standard deviation 0.33. However, during reconstruction, the blur estimate was assumed to be uniform of size 3×3 pixels. Figure 6(c) shows the result of the proposed method for this case. The quality of the reconstructed image is again good.

Next, we demonstrate that the proposed method can be used effectively in real situations also. From the well-known mobile calendar video sequence, four consecutive frames of dimension 160×350 pixels having pure translational motion were taken and downsampled by a factor of 2 to obtain the LR observations. Figure 7(a) shows one such frame. Motion estimates were obtained using the technique proposed by Peleg and Irani.[15] The defocus blur was approximated as Gaussian. The motion blur was neglected as it was small. The output of the proposed algo-

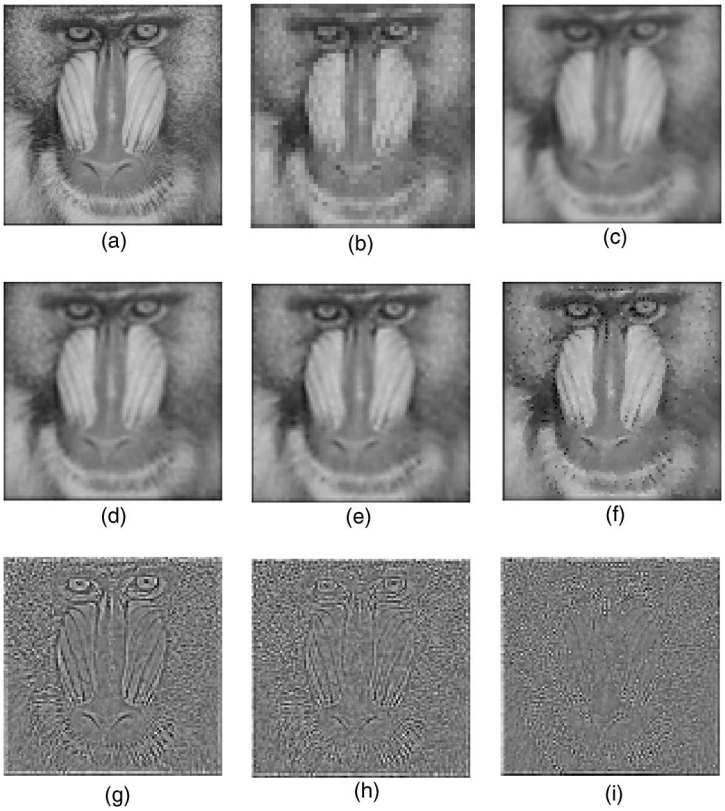

Fig. 5. (a) Original image. (b) One of the LR observations. Results corresponding to (c) bilinear interpolation, (d) GMRF, (e) HMRF, and (f) DAMRF. Error image corresponding to (g) GMRF, (h) HMRF, and (i) DAMRF.

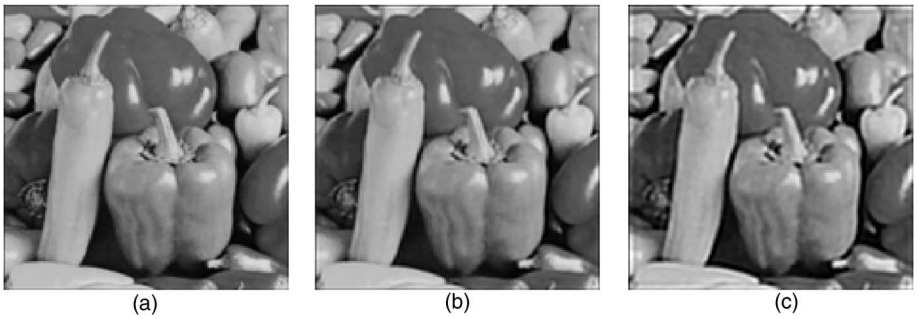

Fig. 6. (a) Original image. (b) Output with motion estimation error. (c) Output with error in blur estimation.

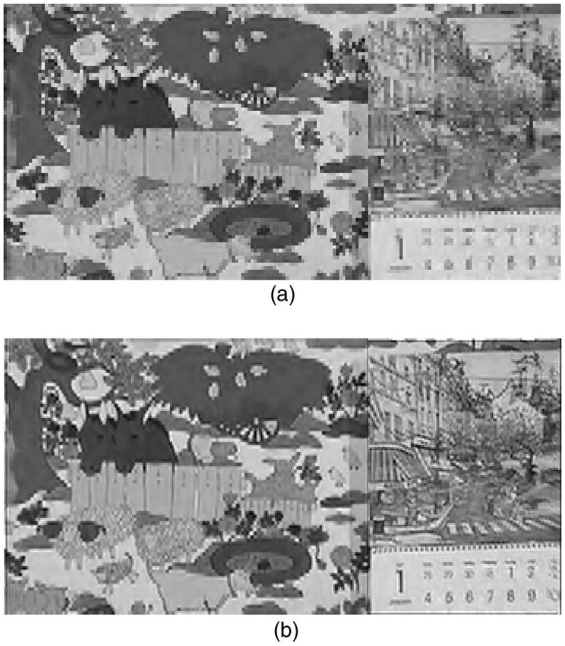

Fig. 7. Results for the mobile calendar sequence. (a) LR observation and (b) DAMRF output.

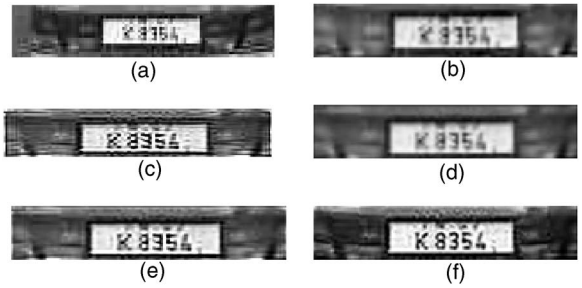

Fig. 8. Results for a real traffic video. (a) Cropped license plate and (b) bilinear interpolation result. Superresolved image using (c) LS, (d) GMRF, (e) HMRF, and (f) DAMRF.

rithm using four LR frames is shown in Fig. 7(b). There is considerable improvement in the overall quality of the image. The numerals become readable in the calendar after superresolution.

Finally, we considered a difficult example from a real traffic video sequence to test the robustness of the proposed method in comparison with other methods. The objective was to superresolve the license plate number of a moving car, given LR frames captured by a handy cam from a height of about 20 feet. The license plate of the car from the LR frames was cropped manually, and one of these frames is shown in Fig. 8(a). Note that it is quite difficult to decipher the number from the LR frame. Since the vehicle was moving away from the camera, the motion blur was predominantly in the vertical direction. The motion estimates were computed using the method in Ref. 15. The camera defocus blur was modeled as Gaussian. Note that the motion as well as the blur estimates are known only approximately here. Four consecutive frames of the scene were used to perform superresolution using different techniques. The reconstructed image using the LS technique [Fig. 8(c)] is poor, as it is very sensitive to errors in motion and blur estimates. The result of the GMRF algorithm shown in Fig. 8(d) is quite blurred. HMRF performs relatively better [Fig. 8(e)], but some of the numbers on the number plate are not easily readable. For example, the second digit "3" can be confused with "9," while the last digit "4" can be misinterpreted as "6." In comparison, the proposed DAMRF algorithm yields the best result, with distinctly defined edges as shown in Fig. 8(f). All the numbers and letters can be read clearly without any ambiguity.

6. CONCLUSIONS

In this paper, we proposed a superresolution algorithm based on a DAMRF prior model. The DAMRF model was shown to be robust to errors in motion and blur estimates while preserving edges. We derived the exact posterior neighborhood structure in the presence of warping, blurring, and downsampling operations. The locality property of the posterior distribution was effectively utilized to reduce the computations significantly. The proposed algorithm was tested on several synthetic and real images, and its performance was found to be very good.

ACKNOWLEDGMENT

The authors thank the reviewers for their constructive and useful comments, which helped in improving the presentation of our work.

Corresponding author K. V. Suresh can be reached by e-mail at sureshkvsit@yahoo.com, A. N. Rajagopalan at raju@ee.iitm.ac.in.

REFERENCES

1. S. Borman and R. L. Stevenson, "Super-resolution from image sequences—a review," in *Proceedings of 1998 Midwest Symposium on Circuits and Systems* (Institute of Electrical and Electronics Engineers, 1998), pp. 374–378.
2. S. Chaudhuri, *Super-Resolution Imaging* (Kluwer, 2001).
3. M. G. Kang and S. Chaudhuri, "Super-resolution image reconstruction," IEEE Signal Process. Mag. **20**, 21–36 (2003).
4. S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," Int. J. Imaging Syst. Technol. **14**, 47–57 (2004).
5. M. Elad and A. Feuer, "Restoration of a single super-resolution image from several blurred, noisy and under-sampled measured images," IEEE Trans. Image Process. **6**, 1646–1658 (1997).
6. D. Rajan and S. Chaudhuri, "Generation of super-resolution images from blurred observations using Markov random fields," in *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, 2001), pp. 1837–1840.
7. A. N. Rajagopalan and V. P. Kiran, "Motion-free super-resolution and the role of relative blur," J. Opt. Soc. Am. A **20**, 2022–2032 (2003).
8. N. K. Bose, M. K. Ng, and A. C. Yau, "A fast algorithm for image super-resolution from blurred observations," EURASIP J. Appl. Signal Process **2006**, 35726 (2006).
9. S. Baker and T. Kanade, "Limits on super-resolution and how to break them," IEEE Trans. Pattern Anal. Mach. Intell. **24**, 1167–1183 (2002).
10. Q. Wang, X. Tang, and H. Shum, "Patch-based blind super-resolution," in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, 2005), pp. 709–716.
11. T. S. Huang and R. Y. Tsai, "Multiframe image restoration and registration," in *Advances in Computer Vision and Image Processing*, T. S. Huang, ed. (JAI Press, 1984), Vol. 1, pp. 317–339.
12. S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy under-sampled multiframes," IEEE Trans. Acoust., Speech, Signal Process. **38**, 1013–1027 (1990).
13. H. Ur and D. Gross, "Improved resolution from sub-pixel shifted pictures," CVGIP: Graph. Models Image Process. **54**, 181–186 (1992).
14. D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacements," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, 1988), pp. 742–746.
15. S. Peleg and M. Irani, "Improving resolution by image registration," CVGIP: Graph. Models Image Process. **53**, 231–239 (1991).
16. R. L. Stevenson and R. R. Schultz, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process. **5**, 996–1011 (1996).
17. R. C. Hardie, K. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of under-sampled images," IEEE Trans. Image Process. **6**, 1621–1632 (1997).
18. A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, 2001), pp. 645–650.
19. E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate sub-pixel registration," IEEE Trans. Image Process. **12**, 826–837 (2003).
20. S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," IEEE Trans. Image Process. **13**, 1327–1344 (2004).
21. S. Z. Li, *Markov Random Field Modeling in Computer Vision* (Springer-Verlag, 1995).
22. J. Besag, "On the statistical analysis of dirty pictures," J. R. Stat. Soc. Ser. B (Methodol.) **48**, 259–302 (1986).
23. S. Chaudhuri and A. N. Rajagopalan, *Depth from Defocus: A Real Aperture Imaging Approach* (Springer-Verlag, 1999).
24. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," IEEE Trans. Image Process. **6**, 721–741 (1984).
25. A. Zomet and S. Peleg, "Super-resolution from multiple images having arbitrary mutual motion," in *Super-Resolution Imaging*, S. Chaudhuri, ed. (Kluwer, 2001), pp. 195–209.