Superresolution Image Reconstruction Using Fast Inpainting Algorithms

Tony F. Chan* Michael K. Ng† Andy C. Yau‡ Andy M. Yip§

September 19, 2006

Abstract

The main aim of this paper is to employ the total variation (TV) inpainting model to superresolution imaging problems. We focus on the problem of reconstructing a high-resolution image from several decimated, blurred and noisy low-resolution versions of the high-resolution image. We propose a general framework for multiple shifted and multiple blurred low-resolution image frames which subsumes several well-known superresolution models. Moreover, our framework allows an arbitrary pattern of missing pixels and in particular missing frames. The proposed model combines the TV inpainting model with the framework to formulate the superresolution image reconstruction problem as an optimization problem. A distinct feature of our model is that in regions without missing pixels, the reconstruction process is regularized by TV minimization whereas in regions with missing pixels or missing frames, they are reconstructed automatically by means of TV inpainting. A fast algorithm based on fixed-point iterations and preconditioning techniques is investigated to solve the associated Euler-Lagrange equations. Experimental results are given to show that the proposed TV superresolution imaging model is effective and the proposed algorithm is efficient.

1 Introduction

Image superresolution refers to a process that increases spatial resolution by fusing information from a sequence of images (with partial overlap in successive elements or frames in, for example, video), acquired in one or more of several possible ways. For brevity, in this context, either the term *superresolution* or *high-resolution* is used to refer to any algorithm which produces an increase in resolution from *multiple low-resolution* degraded images. At least two nonidentical images are required to construct a higher resolution version. The low-resolution frames may be displaced with respect to a reference frame (e.g. LANDSAT images, where there is a considerable

*Department of Mathematics, University of California, Los Angeles, CA 90034-1555, USA. E-mail: chan@math.ucla.edu.

†Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong. Research supported in part by Hong Kong Research Grants Council Grant Nos. 7046/03P, 7035/04P and 7035/05P and HKBU FRGs. E-mail: mng@math.hkbu.edu.hk.

‡Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong.

§Department of Mathematics, National University of Singapore, 2, Science Drive 2, Singapore 117543, Singapore. Research supported in part by National University of Singapore Research Grants R-146-050-079-133 and R-146-050-079-101. E-mail: matymha@nus.edu.sg.

distance between camera and scene), blurred (due to causes like optical aberration, relative motion between camera and object, atmospheric turbulence), rotated and scaled (due to video camera motion like zooming, panning, tilting).

Due to hardware cost, size, and fabrication complexity limitations, imaging systems like charge coupled device (CCD) detector arrays often provide only multiple low-resolution degraded images. However, a high-resolution image is indispensable in applications including health diagnosis and monitoring, military surveillance, and terrain mapping by remote sensing. Other intriguing applications for high-resolution imaging include substituting expensive high-resolution instruments like scanning electron microscopes by their cruder, cheaper counterparts and then applying technical methods to increase the resolution to that derivable with much more costly equipment. Resolution improvement by applying tools from digital signal processing technique has, therefore, been a topic of very great interest.

On the other hand, image inpainting refers to the fill-in of missing or occluded regions in an image based on information available on the observed regions. A common principle for inpainting is to complete isophotes (level sets) in a natural way. Several successful inpainting models have been proposed such as Masnou and Morel [31], Bertalmio *et al.* [3], Chan and Shen [12] and the references therein for other recent models. Chan and Shen proposed in [12] a TV inpainting model which uses variational methods in inpainting. The basic ingredient is to solve the boundary value problem:

$$\min_u \int_\Omega |\nabla u| \quad \text{subject to} \quad u = u_0 \quad \text{in } \Omega \setminus \mathcal{E}.$$

Here $\mathcal{E}$ is the missing region to be inpainted, u_0 is the observed image whose value in $\mathcal{E}$ is missing. Thus the TV inpainting method simply fills in the missing region such that the TV in Ω is minimized. The use of TV-norm is desirable because it has the effect of extending level sets into $\mathcal{E}$ without smearing discontinuities along the tangential direction of the boundary of $\mathcal{E}$. To make the reconstruction robust to noise near the boundary of $\mathcal{E}$ and to denoise the image simultaneously, the constraint $u = u_0$ is relaxed to a regularized least squares fitting, c.f. (6).

In [13], Chan *et al.* considered the problem of inpainting blurry images with unknown point spread functions. In this paper, TV inpainting is also used to fill in missing pixels. They showed that the degradation by missing pixels and by blurring must be recovered simultaneously in order to obtain good quality results. Thus, it is essential that one needs a mechanism which increases the resolution of the image and fills in the missing pixels and/or missing frames simultaneously.

The main aim of this paper is to employ the total variation (TV) inpainting model to superresolution imaging problems. We focus on the problem of reconstructing a high-resolution image from several decimated, blurred and noisy low-resolution versions of the high-resolution image. We propose a general framework for multiple shifted and multiple blurred low-resolution image frames which subsumes several well-known superresolution models. Moreover, our framework allows an arbitrary pattern of missing pixels and in particular missing frames. The proposed model combines the TV inpainting model with the framework to formulate the superresolution image reconstruction problem as an optimization problem. A distinct feature of our model is that in regions without missing pixels, the reconstruction process is regularized by TV minimization whereas in regions with missing pixels or missing frames, they are reconstructed automatically by means of TV inpainting. A fast algorithm based on fixed-point iterations and preconditioning techniques is investigated to solve the associated Euler-Lagrange equations. Experimental results are given to show that the proposed TV superresolution imaging model is effective and

the proposed algorithm is efficient.

The outline of the paper is as follows. In §2, we review the image superresolution reconstruction technique, and present a mathematical formulation of the superresolution imaging problem. In §3, we present our TV inpainting model for the superresolution imaging problem. In §4, we discuss a fast numerical solver for our problem. Finally, numerical results and concluding remarks are given in §5.

2 Review of Superresolution Image Reconstruction

Different methods to obtain superresolution images include the non-uniform interpolation approach, frequency domain approach, and regularization based reconstruction techniques which may be either deterministic or stochastic. A comprehensive survey on superresolution imaging can be found in [14, 44]. For analysis, algorithms and applications of superresolution imaging, see the three recent special issues [34, 36, 38].

The superresolution idea was first proposed by Tsai and Huang [57]. They used a frequency domain approach to demonstrate the ability to reconstruct a single improved resolution image from several down-sampled, noise-free versions of it. Kim, Bose and Valenzuela [22] proposed a weighted recursive least squares algorithm based on sequential estimation theory for filtering and interpolating in the Fourier transform or wavenumber domain. Their objective is to construct a high-resolution image from a registered sequence of undersampled, noisy and blurred frames, displaced horizontally and vertically from each other (sufficient for LANDSAT type imaging). The attainment of image superresolution was based on the feasibility of reconstruction of two-dimensional bandlimited signals from nonuniform samples [23] arising from frames generated by microscanning, i.e., subpixel shifts between successive frames, each of which provides a unique snapshot of a stationary scene. Kim and Su [24] incorporated explicitly the deblurring computation into the high resolution image reconstruction process since separate deblurring of input frames would introduce the undesirable phase and high wavenumber distortions of those frames. A discrete cosine transform-based approach in the spatial domain with regularization, but without the recursive updating feature, was considered in [47]. The most recent activities following the paper published in [22] in this vibrant area have been summarized in some typical papers [46] (galactical image, X-ray image, satellite image of hurricane, city aerial image, CAT-scan of thoracic cavity), [25] (digital electron microscopy), [45] (super-resolution in magnetic resonance imaging) that serve to offer credence to the immense scope, diversity of applications, and the importance of the subject-matter.

Wavelet methods offer considerable promise in the fast interpolation of unevenly spaced data. Motivated by the promise of wavelets, a couple of papers on wavelet superresolution have appeared [43, 7, 28, 9, 10]. These papers use the first generation wavelets and do not subscribe to the need for selecting the mother wavelet to optimize performance. For the selection of the mother wavelet in superresolution imaging problems, see [39].

A minimum mean squared error approach for multiple image restoration, followed by interpolation of the restored images into a single high-resolution image has been presented in [51]. Ur and Gross [58] used the Papoulis and the Brown generalized sampling theorem to obtain an improved resolution image from a set of spatially shifted observations. In these papers, these shifts are assumed to be known. Bose *et al.* adapted a recursive total least squares (TLS) algorithm to tackle high resolution reconstruction from low-resolution noisy sequences with displacement

error during image registration [5]. A theory was advanced, through variance analysis, to assess the robustness of this TLS algorithm for image reconstruction [6]. Specifically, it was shown that under appropriate assumptions the image reconstructed using the TLS algorithm has minimum variance with respect to all unbiased estimates.

A different approach towards superresolution from that in [22] was suggested by Irani and Peleg [21], who used a rigid model instead of a translational model in the image registration process and then applied the iterative back-projection technique from computer-aided tomography. A summary of these and other research during the last decade is contained in a recent paper [16]. Mann and Picard [30] proposed the projective model in image registration because their images were acquired with a video camera. The projective model was subsequently used by Lertrattanapanich and Bose [27] for videomosaicing and high-resolution image reconstruction.

A set theoretic approach to the superresolution restoration problem was suggested by Tekalp *et al.* [55]. They defined convex sets which represent tight constraints on the image to be restored. The projections onto convex sets (POCS) based approach describes a method to incorporating the prior knowledge about the high-resolution image into the reconstruction process. Sauer and Allebach [50] applied the POCS algorithm to this problem subject to the blur-free assumption. Stark and Oskoui [52] applied POCS in the blurred but noise-free case. Tekalp, Ozkan and Sezan [56] formulated a POCS algorithm to compute an estimate from low-resolution images obtained by either scanning or rotating an image with respect to the CCD image acquisition sensor array or mounting the image on a moving platform [20]. Elad and Feuer [15] proposed a unified methodology for superresolution restoration from several geometrically wrapped, blurred, noisy and down-sampled observations by combining maximum likelihood, maximum a posteriori (MAP) and POCS approaches.

Ng *et al.* [35] developed a regularized constrained total least squares solution to obtain a high-resolution image. They considered the presence of perturbation errors of displacements around the ideal subpixel locations in addition to sensor noise. The superiority of the approach over conventional least squares based approach is substantiated through examples. An analysis of the effect of displacement errors on the convergence rate of the iterative approach for solving the transform based preconditioned system of equations during high resolution image reconstruction with multiple sensors has been carried out in [32, 33].

2.1 The Imaging Model

An image acquisition system composed of an array of sensors, where each sensor has a subarray of sensing elements of suitable size, has recently been popular for increasing the spatial resolution with high signal-to-noise ratio beyond the performance bound of technologies that constrain the manufacture of imaging devices. The technique for reconstructing a high-resolution image from data acquired by a prefabricated array of multisensors was advanced by Bose and Boo [4]. In this paper, we focus on the problem of reconstructing a high-resolution image from several blurred low-resolution image frames. The image frames consist of decimated, blurred and noisy versions of the high-resolution image [1, 48].

Let $u[i, j]$ be the original high-resolution image defined on the lattice

$$\Omega = \{1, \ldots, M\} \times \{1, \ldots, N\}.$$

Suppose we have $L \times L$ low-resolution frames $f_{p,q}$ for $p, q = 1, \ldots, L$. The (p, q)-th low-resolution frame is shifted horizontally and vertically by p and q high-resolution pixels respectively with

respect to the $(1,1)$-th reference frame. Each low-resolution frame has size $m \times n$ (with $M = mL$ and $N = nL$), obtained by

$$f_{p,q}[k,l] = \frac{1}{L^2} \sum_{t=-\left[\frac{L-1}{2}\right]}^{\left[\frac{L-1}{2}\right]} \sum_{s=-\left[\frac{L-1}{2}\right]}^{\left[\frac{L-1}{2}\right]} w_{s,t} u[(k-1)L+p+s, (l-1)L+q+t]$$

for $k = 1, \ldots, m$, $l = 1, \ldots, n$. Here $[x]$ is the smallest integer being greater than or equal to x, and

$$w_{s,t} = \begin{cases} 1, & -\left[\frac{L-1}{2}\right] + 1 \le s,t \le \left[\frac{L-1}{2}\right] - 1, \\ \frac{1}{4}, & |s| = |t| = \left[\frac{L-1}{2}\right] \\ \frac{1}{2}, & \text{otherwise.} \end{cases}$$

A pixel intensity value in the low-resolution frame comes from exactly L^2 pixel intensity values in the high-resolution frame. When L is even, the low-resolution frames are shifted relative to each other by the half-pixel value. In addition, each low-resolution frame involves the pixel values outside the scene $u[\cdot,\cdot]$. In our study, the pixel values outside the scene will be determined by the reflective boundary condition, see [37, 42] for details.

In matrix-vector notation, we have

$$\mathbf{f}_{p,q} = D_{p,q} C \mathbf{u}$$

where C is an MN-by-MN averaging matrix and $D_{p,q}$ is an mn-by-MN down-sampling matrix corresponding to the (p,q)-th low-resolution frame, and $\mathbf{u}$ is the MN-by-1 lexicographically ordered vector containing the intensity values of the pixels in the high-resolution image. More precisely, we have

$$C = C^x \otimes C^y$$

where C^x and C^y are N-by-N and M-by-M matrices respectively given by the form

$$C^{x,y} = \frac{1}{L} \begin{pmatrix} 1 & \cdots & \cdots & 1 & & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ 1 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 1 \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & & 1 & \cdots & \cdots & 1 \end{pmatrix} + \frac{1}{L} \begin{pmatrix} 1 & \cdots & 1 & & & & & \\ \vdots & \iddots & & & & & & \\ 1 & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & 1 \\ & & & & & & \iddots & \vdots \\ & & & & & 1 & \cdots & 1 \end{pmatrix}.$$

if L is an odd number;

$$C^{x,y} = \frac{1}{L}\begin{pmatrix} 1 & \cdots & 1 & \frac{1}{2} & & & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & & & \\ 1 & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \frac{1}{2} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \frac{1}{2} \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & 1 \\ & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & & \frac{1}{2} & 1 & \cdots & 1 \end{pmatrix} + \frac{1}{L}\begin{pmatrix} 1 & \cdots & 1 & \frac{1}{2} & & & & \\ \vdots & \iddots & \iddots & & & & & \\ 1 & \iddots & & & & & & \\ \frac{1}{2} & & & & & & & \\ & & & & & & & \frac{1}{2} \\ & & & & & & \iddots & 1 \\ & & & & & \iddots & \iddots & \vdots \\ & & & & \frac{1}{2} & 1 & \cdots & 1 \end{pmatrix}$$

if L is an even number. Note that when L is odd, the bandwidth of C^x and C^y is exactly L whereas when L is even, their bandwidth is $L+1$. The second matrix in $C^{x,y}$ is due to the reflective boundary conditions imposed for pixels outside the scene. The down-sampling matrix $D_{p,q}$ corresponding to the (p,q)-th low-resolution frame is given by

$$D_{p,q} = D_p^x \otimes D_q^y$$

where D_p^x is an n-by-nL down-sampling matrix whose entries are zero except for the $(k, p+(k-1)L)$-th entry for $k = 1, 2, \cdots, n$ which is equal to 1. The matrix D_q^y is an m-by-mL down-sampling matrix defined similarly.

The low-resolution frames are further degraded by a blurring kernel $h_{p,q}$ of window size $W \times W$ and an additive noise $\eta_{p,q}$, yielding

$$g_{p,q}[k,l] = \sum_{t=-\frac{W+1}{2}}^{\frac{W+1}{2}} \sum_{s=-\frac{W+1}{2}}^{\frac{W+1}{2}} h_{p,q}[s,t] f_{p,q}[k+s, l+t] + \eta_{p,q}[k,l].$$

For simplicity, we assume that W is an odd number. The above equation can be rewritten as

$$\mathbf{g}_{p,q} = H_{p,q}\mathbf{f}_{p,q} + \mathbf{n}_{p,q} = H_{p,q}D_{p,q}C\mathbf{u} + \mathbf{n}_{p,q}$$

where $H_{p,q}$ is an mn-by-mn blurring matrix. By interlacing the low-resolution frames, we obtain the following image observation equations:

$$\mathbf{g} = \sum_{p,q} D_{p,q}^T \mathbf{g}_{p,q} = \sum_{p,q} D_{p,q}^T (H_{p,q}D_{p,q}C\mathbf{u} + \mathbf{n}_{p,q}) := A\mathbf{u} + \mathbf{n} \tag{1}$$

where

$$A = \sum_{p,q} D_{p,q}^T H_{p,q} D_{p,q} C \quad \text{and} \quad \mathbf{n} = \sum_{p,q} D_{p,q}^T \mathbf{n}_{p,q}.$$

2.2 Missing Low-resolution Frames

In the above image observation model (1), there are no missing values. Let $\mathcal{D} \subset \Omega$ be a subset of pixels at which the value of $\mathbf{g}$ is observed while pixels in $\Omega \setminus \mathcal{D}$ are missing. Then we have the image observation equation:

$$\begin{aligned} \mathbf{g} &= A\mathbf{u} + \mathbf{n} \\ \mathbf{g}\,|_{\Omega\setminus\mathcal{D}} &\quad \text{missing.} \end{aligned} \tag{2}$$

Let $\Lambda_{\mathcal{D}}$ be the matrix which samples pixels in $\mathcal{D}$ from Ω. Then the image observation equation is given by

$$\mathbf{g}_{\text{obs}} = \Lambda_{\mathcal{D}}(A\mathbf{u} + \mathbf{n}). \tag{3}$$

2.3 Multiple Blurred Images

In addition to the displacement-based superresolution, the motion-free superresolution can be employed to obtain the spatial enhancement. Rajan and Sabhasis [48] considered an availability of decimated, blurred and noisy versions of a high-resolution image which are used to generate a super-resolved image. A known blur (e.g., Gaussian blurs of different values of variances) is used to generate high-resolution images. Suppose the same scene is observed R times but each time the blurring matrices $H_{p,q,r}$ are different, the image observation equations become

$$\mathbf{g}_{\text{obs},r} = \Lambda_{\mathcal{D}_r}(A_r\mathbf{u} + \eta_r) \qquad r = 1, \ldots, R \tag{4}$$

where

$$A_r = \sum_{p,q} D_{p,q}^T H_{p,q,r} D_{p,q} C.$$

3 TV Superresoluton Imaging Model

In this section, we propose a unified model for superresolution image reconstruction. Our model allows an arbitrary pattern of missing pixels and generalizes several important existing models.

To invert the degradation process in (4), we propose a Tikhonov-type regularization model which requires minimization of the following energy:

$$E_\lambda(\mathbf{u}) = \frac{1}{2}\sum_{r=1}^{R} \|\Lambda_{\mathcal{D}_r} A_r\mathbf{u} - \mathbf{g}_{\text{obs},r}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}. \tag{5}$$

Here, $\lambda > 0$ is the regularization parameter for controlling the trade-off between the regularity of the reconstructed image and the fidelity to the observe image; $\|\mathbf{u}\|_{\text{TV}}$ is the discrete total variation of $\mathbf{u}$ defined by

$$\|\mathbf{u}\|_{\text{TV}} = \sum_{i=1}^{M}\sum_{j=1}^{N} \sqrt{(u[i+1,j] - u[i,j])^2 + (u[i,j+1] - u[i,j])^2}$$

where the Neumann boundary condition is applied to the pixel values outside $u[i,j]$ for $1 \le i \le M$ and $1 \le j \le N$.

The use of TV regularization for image restoration has been proposed by Rudin, Osher and Fatemi [49]. The distinctive feature of TV regularization is that edges can be preserved. On the other hand, it can be proved in the continuum case that edges will be destroyed when using the more common H^1 regularization. Thus TV regularization is in general more suitable than H^1 regularization for image reconstruction purposes. However, a potential drawback to TV regularization is that it may lead to the so-called stair-casing effect [11]. We refer the reader to [11] for a survey of recent developments on TV image restoration.

In the case of superresolution image reconstruction where information for certain pixels could be missing, the second term in (5) has a dual purpose. Besides controlling the regularity of the

reconstructed image, the TV term also induces an image inpainting mechanism. In [12], Chan and Shen proposed a TV inpainting model for filling-in missing pixels which is given by:

$$\min_{\mathbf{u}} \frac{1}{2}\|\Lambda_{\mathcal{D}}\mathbf{u} - \mathbf{u}_{\text{obs}}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}. \tag{6}$$

Here, $\mathbf{u}_{\text{obs}}$ is an observed image with missing pixels, $\mathcal{D}$ is the set of observed pixels and $\Lambda_{\mathcal{D}}$ is a down-sampling matrix from the full $M \times N$ grid to $\mathcal{D}$. The main idea of this model is to fill-in the missing pixels in such a way that the overall total variation of the image is minimized. Roughly speaking, the TV will be minimized when the missing pixels are filled-in by a piecewise constant extension into the missing region. In this way, edges present at the interface between the missing and observed regions can be extended inside the missing region. Thus edges can be reconstructed. However, isophotes may not be extended smoothly. Figure 1 shows a particular case where the resolution is increased by 3 folds and the off diagonal frames are missing. Our framework using TV inpainting model can handle arbitrary missing frames and missing pixels within each frame.

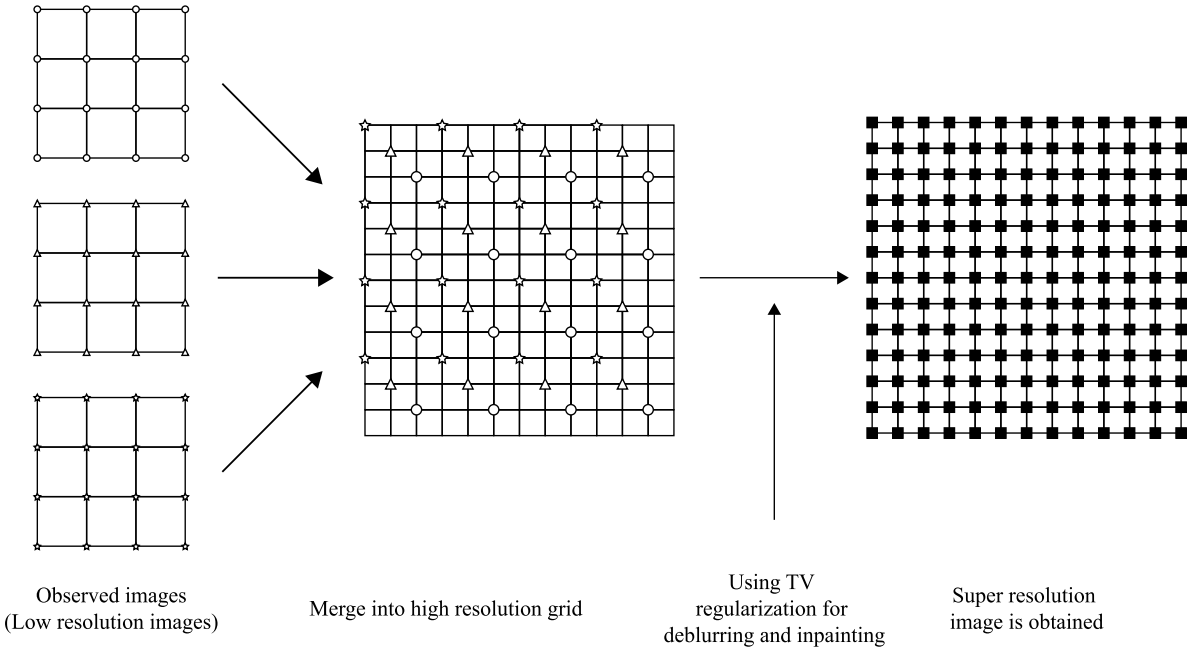


Figure 1: An example using TV inpainting model for the image reconstruction where the resolution is increased by 3 folds and the off diagonal frames are missing.

In the framework of multi-frame and multi-blur superresolution image reconstruction where the low-resolution pixels are modeled as averages of the high-resolution ones, if there are missing pixels in the low-resolution frames, then there may be no information available for some high-resolution pixels. In this case, the high-resolution pixels can only be reconstructed by means of inpainting. In our model, this is done via TV inpainting. In contrast, other high-resolution pixels are reconstructed from the low-resolution ones by reversing the averaging processes with the aid of TV regularization.

Another framework of superresolution image reconstruction is solely based on interpolation where the (only) low-resolution frame is modeled as a down-sampled version of the high-

resolution frame, instead of a down-sampled version of the local average of the high-resolution frame. This framework is also known as *intraframe reconstruction*. Thus, superresolution reconstruction is simply an up-conversion process. In our notation, this framework corresponds to the case where the averaging matrix C is dropped, or more precisely, reset to the identity matrix I. Because of the absence of the local averaging processes, the low-resolution frame encodes no information about the missing high-resolution pixels. Thus, any attempt to reconstruct the missing high-resolution pixels must resort to a form of interpolation, i.e., inpainting. Our model automatically does the inpainting by means of TV minimization.

3.1 Some Special Cases

In this subsection, we give some important special cases of our unified model which have been considered by several authors. However, we use the TV regularization to reconstruct high-resolution images.

3.1.1 Full Set of Low-resolution Frames, Without Blurring

This is the case considered by Bose and Boo [4] and others where a full set of $L \times L$ low-resolution frames are available. Thus in the ideal case where noise and numerical errors are absent, a perfect reconstruction is allowed. In our unified model, this case corresponds to

1. $R = 1$ — no multiple observations for each frame;
2. $\Lambda_{\mathcal{D}_r} = I$ for all r — no missing frames and/or pixels;
3. $H_{p,q,r} = I$ for all p, q, r — no blurring in each frame.

By dropping the subscript r, the model (5) is then simplified to:

$$E_\lambda(\mathbf{u}) = \frac{1}{2}\|\tilde{A}\mathbf{u} - \mathbf{g}_{\text{obs}}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}$$

where

$$\tilde{A} = \Lambda_{\mathcal{D}_r} A = \sum_{p,q} D_{p,q}^T D_{p,q} C = C.$$

We remark that for the sake of simplicity, we here ignore the sub-pixel displacement error in each frame [4]. To incorporate the frame-specific displacement errors, we simply replace the definition of $\tilde{A}$ by

$$\tilde{A} = \sum_{p,q} D_{p,q}^T D_{p,q} C_{p,q}$$

where $C_{p,q}$ is a shifted averaging matrix with shift $(\epsilon_{p,q}^x, \epsilon_{p,q}^y)$, see [4].

3.1.2 Partial Set of Low-resolution Frames, Without Blurring

This case is similar to the first case above, except that some of the low-resolution frames are missing. This setting has been considered by Nguyen and Milanfar [43] and other authors. Let $F = \{(p, q) :\ 1 \le p \le L,\ 1 \le q \le L\}$ be the full set of low-resolution frames. Let $F' \subset F$ be a partial set of frames there are observed. Then, the set of observed pixels is given by

$$\mathcal{D}_r = \{(p + (k-1)L, q + (l-1)L) :\ k = 1, 2, \ldots, m,\ l = 1, 2, \ldots, n,\ (p, q) \in F'\}.$$

Our model reduces to this case when

1. $R = 1$ — no multiple observations for each frame;
2. $\Lambda_{\mathcal{D}_r}$ is the down-sampling matrix which samples $\mathcal{D}_r$ from Ω for all r — only a subset F' of frames is observed;
3. $H_{p,q,r} = I$ for all p, q, r — no blurring in each frame.

Then, the model (5) can be simplified to

$$E_\lambda(\mathbf{u}) = \frac{1}{2}\|\tilde{A}\mathbf{u} - \mathbf{g}_{\text{obs}}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}$$

where

$$\tilde{A} = \Lambda_{\mathcal{D}_r} A = \Lambda_{\mathcal{D}_r} C.$$

3.1.3 Single Low-resolution Frame, With Multiple Blurring

This setting is considered by Rajan and Sabhasis [48]. An interpretation of this setting is that when shifted low-resolution frames are unavailable, one can still gather more information about the scene by acquiring R different blurred versions of the same scene. This is covered by our model with

1. $R > 1$ — with multiple observations for each frame;
2. $\Lambda_{\mathcal{D}_r} = D_{1,1}$ for all r — no shifted frame is observed.

The model (5) becomes

$$E_\lambda(\mathbf{u}) = \frac{1}{2}\sum_{r=1}^{R} \|\tilde{A}_r\mathbf{u} - \mathbf{g}_{\text{obs},r}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}$$

where

$$\tilde{A}_r = \Lambda_{\mathcal{D}_r} A_r = H_{1,1,r} D_{1,1} C.$$

3.1.4 Intraframe Super-resolution Reconstruction

In this case, the low-resolution frame is simply treated as a down-sampled version of the high-resolution frame. An example application of this setting is digital zooming. To use our model for such a setting, we set

1. $R = 1$ — no multiple observations for each frame;
2. $\Lambda_{\mathcal{D}_r} = D_{1,1}$ for all r — no shifted frame is observed;
3. $H_{p,q,r} = I$ for all p, q, r — no blurring in each frame;
4. $C = I$ — no averaging.

The model (5) now becomes

$$E_\lambda(\mathbf{u}) = \frac{1}{2}\|\tilde{A}\mathbf{u} - \mathbf{g}_{\text{obs}}\|^2 + \lambda\|\mathbf{u}\|_{\text{TV}}$$

where

$$\tilde{A} = D_{1,1} A = D_{1,1}.$$

This model is exactly the TV Inpainting model (6) proposed by Chan and Shen [12] where the missing pixels are filled-in by the TV inpainting process.

3.1.5 Random Missing Pattern

Some sensors on the sensor plate may be damaged for some reason. Therefore, some pixels in the low resolution image will be missed randomly. Our model can be easily adapted to accommodate such a situation — one simply lets $\Lambda_{\mathcal{D}_r}$ be the down-sampling matrix which samples the observed pixels from the full grid.

4 Numerical Solvers

In this section, we discuss how to solve the optimization problem in (5). The Euler-Lagrange equation for the energy is given by the following non-linear system

$$0 = \nabla E_\lambda(\mathbf{u}) = \sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T (\Lambda_{\mathcal{D}_r} A_r \mathbf{u} - \mathbf{g}_{\text{obs},r}) - \lambda L_{\mathbf{u}} \mathbf{u} \tag{7}$$

where $L_{\mathbf{u}}$ is the matrix form of a central difference approximation of the differential operator $\nabla \cdot (\nabla / |\nabla u|)$.

The artificial time marching scheme proposed in Rudin, Osher and Fatemi [49] obtains the solution to (7) as the steady state of a parabolic partial differential equation. Indeed, it is equivalent to employing the following gradient descent method to solve the minimization problem in (5)

$$\mathbf{u}^{i+1} = \mathbf{u}^i - dt \nabla E_\lambda(\mathbf{u}^i) \tag{8}$$

where $dt > 0$ is the time-step parameter restricted by stability conditions (i.e. dt has to be small enough so that the scheme is stable). The drawback of this method is that explicit methods can be very slow due to stability constraints. Implicit methods can also be applied but then one has to deal with the nonlinearity and the solution of the resulting linear systems.

In [59], the lagged diffusivity fixed point iteration is introduced. This method consists of linearizing the nonlinear differential term in (7) by lagging the diffusion coefficient $1/|\nabla u|$ one iteration behind. Thus $\mathbf{u}^{i+1}$ is obtained as the solution to the linear equation:

$$\left[\sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T \Lambda_{\mathcal{D}_r} A_r - \lambda L_{\mathbf{u}^i} \right] \mathbf{u}^{i+1} = \sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T \mathbf{g}_{\text{obs},r}. \tag{9}$$

This algorithm can be interpreted within the framework of generalized Weiszfeld's methods, as introduced in [60]. As proved in that paper, this method is monotonically convergent, in the sense that the objective function evaluated at the iterates form a monotonically decreasing sequence, and that the convergence rate is linear. In practice, this method is very robust.

In this paper, we will employ lagged diffusivity fixed point iteration to solve the TV regularization and inpainting problem. In the next section, we will discuss how to use a good preconditioner for the linear system arising from lagged diffusivity fixed point iteration.

4.1 Preconditioners

The preconditioned conjugate gradient (PCG) method is used to solve (9). Given a matrix equation $B\mathbf{x} = \mathbf{b}$, there are two criteria for choosing a preconditioner for B [18]. First, the preconditioner should be a "good" approximation to B. Secondly, it must be easily inverted.

The matrices we are trying to precondition are the coefficient matrices in (9). The coefficient matrix is the sum of a convolution operator and an elliptic operator. Chan *et al.* [8] proposed a preconditioner for such system by taking the sum of the cosine transform approximations to the convolution operator and the elliptic operator separately. The resulting approximation can still be diagonalized by the discrete cosine transform matrix and is therefore easily invertible. Numerical results in [8] have shown that this approach is very effective. In [37], Ng *et al.* also used cosine transform based preconditioners to precondition the linear system arising from the reconstruction of high-resolution images. When the number of shifted low-resolution images is equal to four in the 2-by-2 sensor setting and these four shifted low-resolution images are shifted relative to each other by the half-pixel value, they showed that the conjugate gradient method, when applied to solving the cosine preconditioned system, converges superlinearly. We note that under the noiseless condition, the four shifted low-resolution images are sufficient to reconstruct the high-resolution image perfectly. In [40], Ng and Sze further modified cosine transform based preconditioners to handle some special cases where the number of shifted low-resolution images is equal to two. Numerical results showed that the performance of these cosine transform based preconditioners are quite good for some special cases. However, the cosine transform based preconditioners do not work well in general.

Our experimental results indicate that the most efficient way to solve the matrix equations arising from high-resolution image reconstruction is to apply the Factorized Sparse Inverse Preconditioner (FSIP) [26, 2, 54] to the system in (9) arising from lagged diffusivity fixed point iteration. In the next section, we will present the numerical results to demonstrate the effectiveness of this preconditioner.

Let B be an n-by-n symmetric positive definite matrix, and let its Cholesky factorization be $B = VV^T$. The idea of FSIP is to find a sparse lower triangular matrix L such that $B^{-1} \approx L^T L$. Let $\mathcal{S}$ be the given sparse pattern such that

$$\{(i,i) \mid i = 1, \cdots, n\} \subset \mathcal{S}$$

and L be such that

$$L(i,j) = 0 \quad \text{if } (i,j) \notin \mathcal{S}.$$

Now we look for L with sparse pattern $\mathcal{S}$ such that the Frobenius norm

$$\|I - VL\|_F$$

is minimized. Kolotilina and Yeremin [26] showed that L can be obtained by the following algorithm:

Algorithm: Construction of FSIP

Step 1. Compute $\hat{L}$ with sparse pattern $\mathcal{S}$ such that

$$[\hat{L}B]_{ij} = \delta_{i,j}, \quad (i,j) \in \mathcal{S}. \tag{10}$$

Step 2. Let $\hat{D} = (\text{diag}(\hat{L}))^{-1}$ and $L = \hat{D}^{\frac{1}{2}}\hat{L}$.

We remark that the computation of L can be done in parallel between rows:

$$\hat{L}(i, \mathcal{S}_i)B(\mathcal{S}_i, \mathcal{S}_i) = [0, \ldots, 0, 1],$$

where $\mathcal{S}_i = \{j|(i,j) \in \mathcal{S}\}$ and $B(\mathcal{I}, \mathcal{J})$ denotes the submatrix of B containing the rows with index set $\mathcal{I}$ and columns with index set $\mathcal{J}$. The preconditioning step has a natural parallelism since the multiplication of $L^T L$ to a residual vector r can be computed in parallel.

In [29], Lin *et al.* considered the Factorized Banded Inverse Preconditioner (FBIP), a special type of FSIP. They set sparse pattern to be banded, i.e.,

$$\mathcal{S} = \{(i,j)| \max(1, i-k+1) \leq j \leq i \leq n\},$$

where k is the bandwidth of the factor L. Using (10), the i-th row of $\hat{L}$ can be obtained by solving the linear system

$$\hat{L}(i, i' : i)B(i' : i, i' : i) = [0, \ldots, 0, 1], \quad i = 1, \ldots, n, \tag{11}$$

where $i' = \max(1, i-k+1)$. Recall that B is an n-by-n symmetric positive definite matrix. We note that for each i, the above linear system can be solved in $O(k^3)$ operations, and the cost to obtain the factor L is $O(nk^3)$ operations. When B is a Toeplitz matrix, the construction cost of the FBIP for Toeplitz matrices can be reduced from $O(nk^3)$ to $O(k^2)$. In [29], Lin *et al.* showed that if a Toeplitz matrix T has certain off-diagonal decay property (or T is banded), then the FBIP will be a good approximation to T^{-1}. They also showed that FBIP will be a good approximation to the Toeplitz-related systems of the form $I + T^T DT$, where T is a Toeplitz matrix and D is a positive diagonal matrix.

By solving the TV inpainting model for the reconstruction of high-resolution images, the involved coefficient matrix in (9) is in the form of

$$\begin{bmatrix} B^{(1,1)} & B^{(1,2)} & \ldots & B^{(1,N)} \\ B^{(2,1)} & B^{(2,2)} & \ldots & B^{(2,N)} \\ \vdots & \vdots & \ldots & \vdots \\ B^{(N,1)} & B^{(N,2)} & \ldots & B^{(N,N)} \end{bmatrix}.$$

Here the $M \times M$ blocks $B^{(i,j)}$ are defined by

$$[B^{(i,j)}]_{k,l} = b_{k,l}^{(i,j)}, \quad k, l = 1, 2, \ldots, M.$$

We assume that the blurring matrix has a block-level off-diagonal decay property, and each block also has the off-diagonal decay property. This is true if the blurring function decays in spatial domain, or if the support of the blur is small. In this case, we can set the factor of the factorized inverse preconditioner to be triangular banded block matrix with each block being a

banded matrix. For example,

$$L = \left[\begin{array}{cccc|cccc|cccc|cccc} + & & & & & & & & & & & & & & & \\ + & + & & & & & & & & & & & & & & \\ & + & + & & & & & & & & & & & & & \\ & & + & + & & & & & & & & & & & & \\ \hline + & + & & & + & & & & & & & & & & & \\ + & + & + & & + & + & & & & & & & & & & \\ & + & + & + & & + & + & & & & & & & & & \\ & & + & + & & & + & + & & & & & & & & \\ \hline & & & & + & + & & & + & & & & & & & \\ & & & & + & + & + & & + & + & & & & & & \\ & & & & & + & + & + & & + & + & & & & & \\ & & & & & & + & + & & & + & + & & & & \\ \hline & & & & & & & & + & + & & & + & & & \\ & & & & & & & & + & + & + & & + & + & & \\ & & & & & & & & & + & + & + & & + & + & \\ & & & & & & & & & & + & + & & & + & +\end{array}\right]. \tag{12}$$

Let k and k' be the half bandwidth of each block and the bandwidth of the banded block matrix respectively (for the matrix given by (12), $N = M = 4$ and $k = k' = 2$). The sparse pattern is set to

$$\mathcal{S} = \cup_{i=1}^{MN} \mathcal{S}_i,$$

where $\mathcal{S}_i$ is the index set corresponding to the ith row. Let

$$i = i_0 M + i_1 \quad \text{with } 0 \le i_0 \le N-1 \text{ and } 1 \le i_1 \le M.$$

Let

$$i_0' = \min(i_0, k'-1), \quad i_1' = \max(i_1 - k + 1, 1), \quad \text{and} \quad i_1'' = \min(i_1 + k - 1, M),$$

we define

$$\mathcal{S}_i = \left\{(i,j) | j = i_0 M + i_1' : i\right\} \cup \left(\cup_{i'=1}^{i_0'} \left\{(i,j) | j = (i_0 - i')M + i_1' : (i_0 - i')M + i_1'')\right\}\right).$$

It is easy to see that the number of nonzero entries in each row of L is not greater than

$$k + (k'-1)(2k-1) = 2kk' - k - k' + 1.$$

Thus, the total cost of computing the factor L is bounded by $O(MN(2kk' - k - k' + 1)^3) = O(MNk^3k'^3)$ and the storage requirement is $O(MNkk')$. Similar to the one-dimensional case, if B is a block-Toeplitz-Toeplitz-block (BTTB) matrix, then the factor L is a near BTTB lower triangular matrix and the cost of constructing L only depends on k and k' (independent of M and N), see [29] for details.

Because the construction cost grows cubically with the increase of k and k', the benefit of the decrease in the number of iterations will be offset by the cost of constructing the preconditioner with the increase of k and k'. Therefore, it is important to choose suitable values of k and k' in practical applications. Usually the linear system to be solved is well-conditioned at the first few lagged diffusivity fixed point iterations. Due to the initial cost of constructing the FBIP, it

is too expensive to choose a big k or k' for early iterates. However, as the iterates get close to the solution, the ill-conditioning of the linear system makes choosing bigger k and k' necessary. Therefore, the best strategy is to choose small k and k' at the beginning, and increase k and k' gradually. Here we do not provide a heuristic for selecting k and k', that such heuristics should be investigated in the future, and that k and k' are fixed at 3 in all the numerical experiments in Section 5.

We end this section by giving a summary of the algorithm as follows:

Algorithm: Superresolution Image Reconstruction Using Fast TV Inpainting

Step 0. Set $i = 0$ and input A_r, $\Lambda_{\mathcal{D}_r}$, $\mathbf{g}_{obs,r}$ for $r = 1, 2, \cdots, R$ and an initial guess $\mathbf{u}^0$.
Step 1. Construct the factorized banded inverse preconditioner for the matrix

$$\left[\sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T \Lambda_{\mathcal{D}_r} A_r - \lambda L_{\mathbf{u}^i}\right]$$

Step 2. Solve (9)

$$\left[\sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T \Lambda_{\mathcal{D}_r} A_r - \lambda L_{\mathbf{u}^i}\right] \mathbf{u}^{i+1} = \sum_{r=1}^{R} A_r^T \Lambda_{\mathcal{D}_r}^T \mathbf{g}_{\text{obs},r}$$

by using the preconditioned conjugate gradient method with the factorized banded inverse preconditioner.

Step 3. If $||\mathbf{u}^{i+1} - \mathbf{u}^i||_2$ is less than the given tolerance, then the algorithm stops, otherwise set $i = i + 1$ and goto Step 1.

5 Simulation Results

In the simulation results, we assume that $L = 3$ and the sensor plane samples the image diagonally. Thus, 6 out of the 9 frames are missing. We also assume the blur to be a Gaussian blur which is given by

$$h_{p,q}[s,t] = ce^{-(s^2+t^2)/2\sigma^2}$$

where $c > 0$ is a normalization constant so that $\sum_{s,t} h_{p,q}[s,t] = 1$. The window size of the blurring function is 5-by-5. Figure 2(a) shows the original image "cameraman" and one of the low-resolution frames which are of size 126-by-126 and 42-by-42 respectively. The noise level is measured by SNR (signal-to-noise ratio) which is given by

$$\text{SNR } = 20 \log_{10} \left(\frac{||A\mathbf{u}||_2}{||\mathbf{n}||_2} \right).$$

In our tests, Gaussian noise with SNR of 40 dB is added to the low-resolution images. Such noise level is used, see for instance [4, 9, 48]. For the selection of regularization parameter λ, it is chosen (by trial and error) to minimize the relative error which is given by

$$\frac{\|\mathbf{u} - \mathbf{u}_c(\lambda)\|_2}{\|\mathbf{u}\|_2} \tag{13}$$

where $\mathbf{u}$ is the original image and $\mathbf{u}_c(\lambda)$ is the reconstructed image using the inpainting algorithm with the regularization parameter λ. In the figures and the tables, we report the relative error and also report the PSNR (Peak-signal-to-noise ratio) of the reconstructed image, where it is given by

$$\text{PSNR } = 20 \log_{10} \left(\frac{\max_i [\mathbf{u}_c(\lambda)]_i}{||\mathbf{u} - \mathbf{u}_c(\lambda)||_2} \right).$$

The larger PSNR is, the better the visual quality of the reconstructed image is. In the computation, the parameters k and k' for the factorizaed banded inverse preconditioner are set to 3 in all the examples and the iterations. The results are shown as follows:

- Figure 3 shows the case of multi-frame without blur which is described in Section 3.1.2. The figure shows the 3 observed images and the reconstructed high-resolution (HR) image with $\lambda = 0.0072$.
- Figures 4 and 5 show the case of multi-frame with a single blur. Figure 4 shows the 3 different observed frames blurred by the Gaussian kernel with $\sigma = 0.8$ and the reconstructed HR image with $\lambda = 0.0004$.
- Figure 5 shows the 3 different observed frames blurred by Gaussian kernels with $\sigma = 0.8$, 0.9 and 1.0 respectively from 3 different low resolution frames and the reconstructed HR image with $\lambda = 0.0003$.
- Figure 6 shows the 3 different observed frames blurred by Gaussian kernels with $\sigma = 0.8, 0.9$ and 1.0 respectively from the same low resolution frame and the reconstructed HR image with $\lambda = 0.00005$.
- Figures 7, 8, 9 and 10 show the case of random missing pixels in different frames which is described in the Section 3.1.5. In Figures 7 and 8, the missing pixels from the low-resolution frames are chosen randomly. The probability of damage for each sensor is $\frac{1}{2}$. The percentage of missing pixels of the 3 observed frames are 51.88%, 49.49% and 52.72% for both Figures 7 and 8. Figure 7 shows the observed images without blurring and the reconstructed HR image with $\lambda = 0.0052$.
- Figure 8 shows the observed images blurred by a Gaussian kernel with $\sigma = 0.8$ and the reconstructed HR image with $\lambda = 0.0074$.
- In Figures 9 and 10, the sensor plate is assumed to be damaged by some scratches. Figure 9 shows the observed frames without blurring and the reconstructed image with $\lambda = 0.011$. Figure 10 shows the observed frames with parameter $\sigma = 0.8$ and the reconstructed image with $\lambda = 0.0008$.

The relative errors, computational time, and PSNR of the reconstructed images in Figures 3–10 are summarized in Table 1.

In the tests, we fix the parameters k and k' for the factorized banded inverse preconditioner in the fixed-point iterations, we observe that the proposed algorithm can still reconstruct high-resolution images very quickly (cf. Table 1). In addition, We compared our method and the artificial time marching scheme (8) for TV regularization in Table 3. The artificial time marching scheme is proposed in [49] to solve the Euler-Lagrange equation (7). We note that the

(a)

(b)

Figure 2: (a) The original 126-by-126 'cameraman' image and (b) a 42-by-42 low-resolution frame.

Method	Reg. Para. λ	Rel. Error	CPU Time (s)	PSNR (dB)
Fig. 3: Multi-frame, no blur	0.0072	0.0737	21.67	25.10
Fig. 4: Multi-frame, single blur	0.0004	0.0874	42.92	24.06
Fig. 5: Multi-frame, multi-blur	0.0003	0.0925	95.98	23.62
Fig. 6: Single frame, multi-blur	0.00005	0.128	110.47	22.82
Fig. 7: Random, no blur	0.0052	0.110	31.22	22.91
Fig. 8: Random, single blur	0.0074	0.116	49.66	23.08
Fig. 9: Scratches, no blur	0.011	0.0895	24.25	22.99
Fig. 10: Scratches, single blur	0.0008	0.104	44.61	21.68

Table 1: Summary of the results in Figures 3–10.

computational times given in the table are for the single and fixed value of λ. The results show that our algorithm is much faster than the artificial time marching scheme.

For comparison, Table 2 shows that the comparison of two different regularization functionals. We find that the method with TV regularization in general is better than that with the H^1 regularization where the optimal regularization parameters are used in both regularization methods. Figure 11 shows the reconstructed images of the two regularization functionals. Images in left column are reconstructed by TV regularization while images in right column are reconstructed by H^1 regularization. We observe that images reconstructed using TV regularization are sharper than those using H^1 regularization, especially near the edges. Figure 12 shows the details near the buildings in the background. It can be clearly seen that severe ringing artifacts are present in the images reconstructed by the H^1 method. The checkerboard artifacts resulted from missing frames are gracefully suppressed by the TV reconstruction.

(a)

(b)

(c)

(d)

Figure 3: Multi-frame without blur. 3(a)-3(c) Observed image frames. 3(d) Restored Image with $\lambda = 0.0072$ reconstructed in 21.67 seconds, relative error $= 0.0737$ and PSNR $= 25.10$ dB.

(a)

(b)

(c)

(d)

Figure 4: Multi-frame with a single blur. 4(a)-4(c) Observed image frames with blurring $\sigma = 0.8$. 4(d) Restored image with $\lambda = 0.0004$ reconstructed in 42.92 seconds, relative error = 0.0874 and PSNR = 24.06 dB.

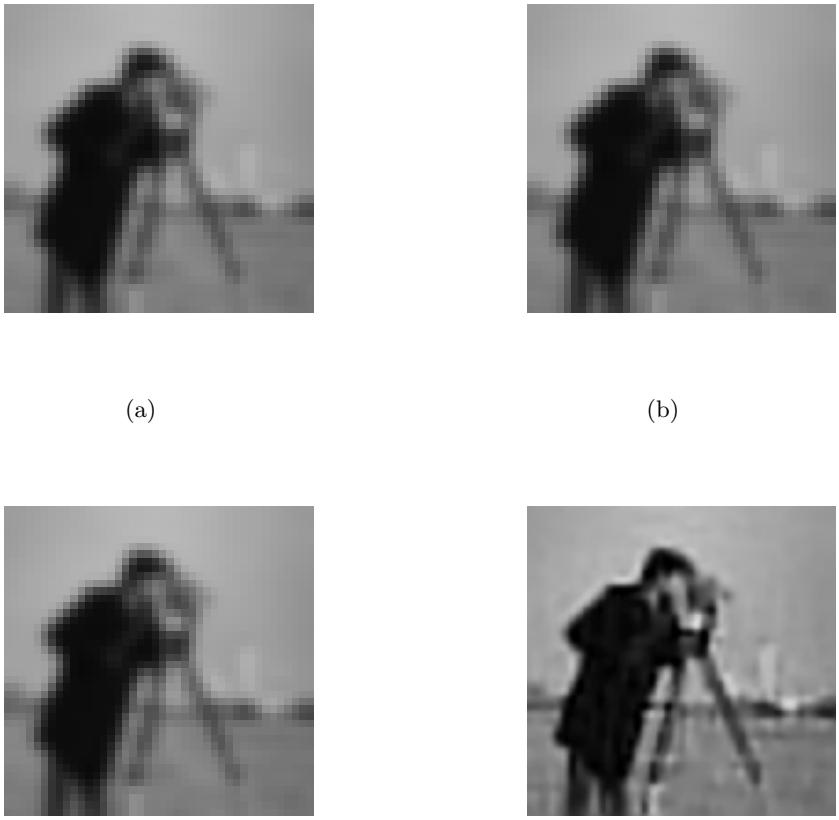
(a)

(b)

(c)

(d)

Figure 5: Multi-frame with multi-blur. 5(a)-5(c) Observed image frames blurred by Gaussian kernels with $\sigma = 0.8$, 0.9 and 1 respectively. 5(d) Restored image with $\lambda = 0.0003$ reconstructed in 95.98 seconds, relative error = 0.0925 and PSNR = 23.62 dB.

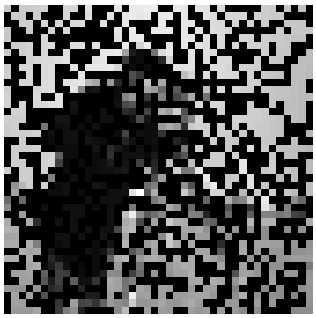


(a) (b)

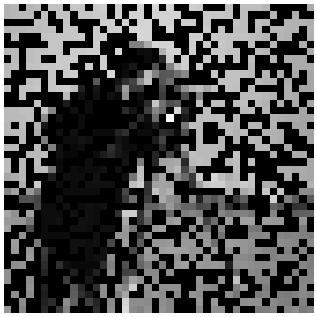
(c) (d)

Figure 6: Single-frame with multi-blur. 6(a)-6(c) Observed image frames blurred by Gaussian kernels with $\sigma = 0.8$, 0.9 and 1 repsectively. 6(d) Restored Image with $\lambda = 0.00005$ reconstructed in 110.47 seconds, relative error = 0.128 and PSNR = 22.82 dB.

(a)



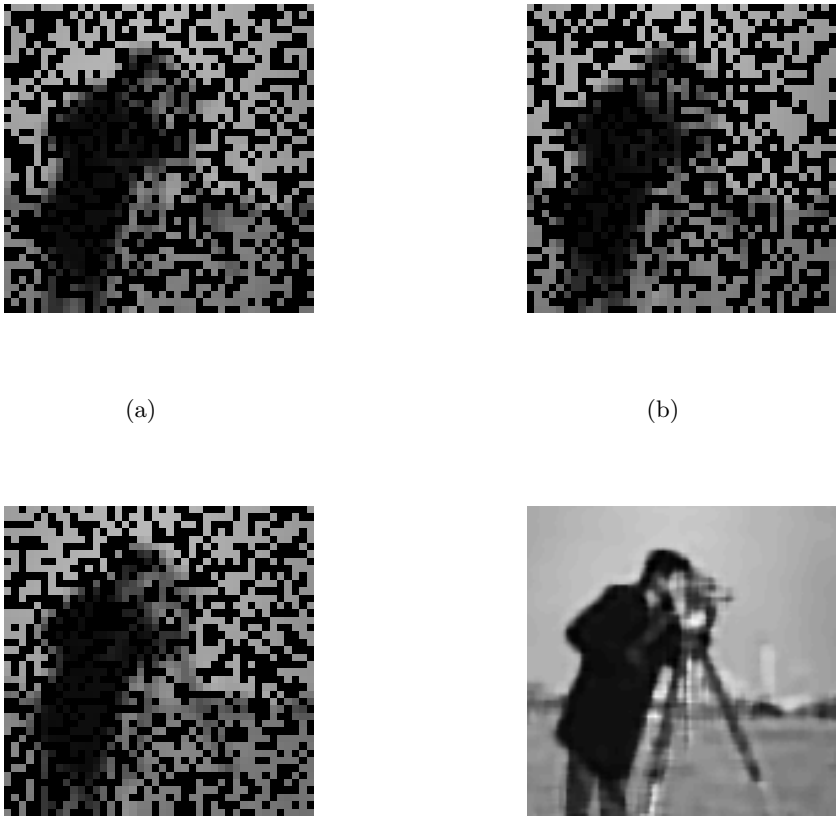
(b)



(c)

(d)

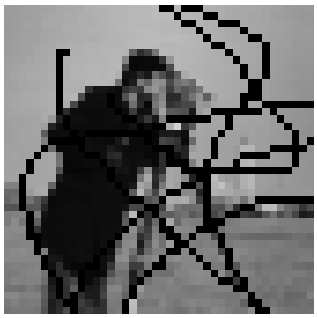
Figure 7: Random missing pixels without blur. 7(a)-7(c) Observed image frames. 7(d) Restored image with $\lambda = 0.0052$ reconstructed in 31.22 seconds, relative error = 0.110 and PSNR = 22.91 dB.



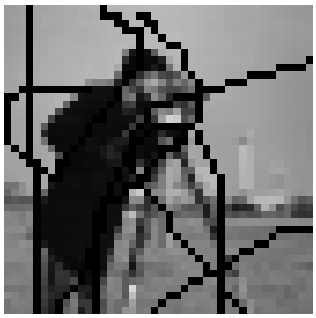
(a) (b)

(c) (d)

Figure 8: Random missing pixels with a single blur. 8(a)-8(c) Observed images frames blurred by a Gaussian kernel with $\sigma = 0.8$, 8(d) Restored image with $\lambda = 0.0074$ reconstructed in 49.66 seconds, relative error $= 0.116$ and PSNR $= 23.08$ dB.



(a)

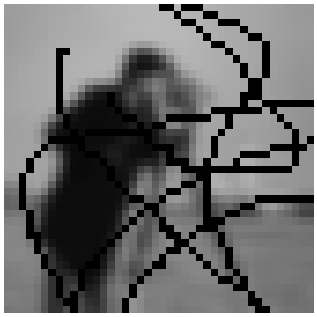


(b)

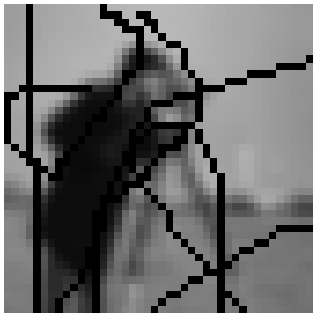
(c)

(d)

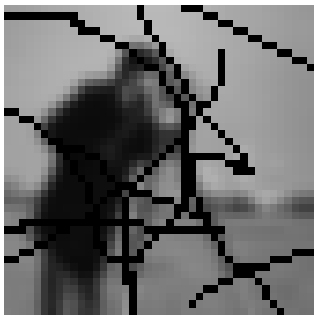
Figure 9: Scratches in the sensor plate without blur.9(a)-9(c) Observed image frames. 9(d) Restored image with $\lambda = 0.011$ reconstructed in 24.25 seconds, relative error = 0.0895 and PSNR = 22.99 dB.



(a)



(b)



(c)

(d)

Figure 10: Scratches in the sensor plate with a single blur. 10(a)-10(c) Observed image frames blurred by a Gaussian kernel with $\sigma = 0.8$. 10(d) Restored image with $\lambda = 0.0008$ reconstructed in 44.61 seconds, relative error = 0.104 and PSNR = 21.68 dB.

6 Concluding Remarks

In this paper, we proposed a general framework for multiple shifted and multiple blurred low-resolution image frames which subsumes several well-known superresolution models. A fast algorithm based on fixed-point iterations and preconditioning techniques is also developed. Our experimental results have shown that the proposed TV superresolution imaging model is quite effective and the proposed algorithm is quite efficient. Here we would like to mention two research issues for further consideration.

1. In this paper, we have not addressed the problem about the selection of the regularization parameter λ. It is obvious that the quality of the reconstruction image and the efficiency of the proposed algorithm will be affected by the selection of the regularization parameter. In the literature, L-curve method [19] and generalized cross-validation method [17] have been used to determine regularization parameters for solving inverse problems. In particular, Ng *et al.* [35] have used L-curve method in constrained total least squares computations for high resolution image reconstruction with multisensors. It is interesting to figure out how to select the regularization parameter efficiently and effectively in superresolution imaging algorithms.

2. Given that the use of the TV term is two-fold, it is used for regularization as well as forcing inpainting. It may be useful to have two TV terms with two different regularization parameters. One is for TV regularization and the other one is for TV inpainting. It is worthwhile to study the performance of such TV norm in superresolution inpainting problems. We note that Strong and Chan [53] have studied weighted TV norm in image processing.

	TV Regularization		H^1 Regularization	
Method	Relative errors	PSNR	Relative errors	PSNR
Multi-frame, no blur	0.0738	25.10 dB	0.0835	23.77 dB
Multi-frame, single blur	0.0874	24.06 dB	0.0900	24.30 dB
Multi-frame, multi-blur	0.0934	24.00 dB	0.0959	24.04 dB
Single frame, multi-blur	0.1280	22.82 dB	0.1310	22.44 dB

Table 2: Comparison with H^1 regularization.

Method	Our method (s)	Artificial time marching scheme (s)
Multi-frame, no blur	21.67	696.64
Multi-frame, single blur	42.92	950.55
Multi-frame, multi-blur	95.98	1870.00
Single-frame, multi-blur	110.47	1826.40

Table 3: Comparison with the artificial time marching scheme.

Acknowledgement:

The authors would like to thank the anonymous reviewers for the helpful suggestions to improve the quality of the manuscript.

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] M. Benzi, C. Meyer, and M. Tuma, A sparse approximate inverse preconditioner for the conjugate gradient method, SIAM J. Sci. Comput., 17 (1996), pp. 1135-1149.

[3] M. Bertalmio, G. Sapiro, V. Caselles, and C. Balleste, Image inpainting, Technical report, 60:259-268, ECE-University of Minnesota, 1999.

[4] N. Bose and K. Boo, High-resolution image reconstruction with multisensors, International Journal of Imaging Systems and Technology, 9 (1998), pp. 294-304.

[5] N. Bose, H. Kim and H. Valenzuela, Recursive total least squares algorithm for image reconstruction from noisy undersampled frames, Multidimensional Systems and Signal Processing, 4 (1993), pp. 253-268.

[6] N. Bose, H. Kim and B. Zhou, Performance analysis of the TLS algorithm for image reconstruction from a sequence of undersampled noisy and blurred frames, International Conference on Image Processing (ICIP '94), Austin, Texas, November 13-16, 1994, pp. 571-575.

[7] N. Bose and S. Lertrattanapanich, Advances in wavelet superresolution, Proceedings of International Conference on Sampling Theory and Application SAMPTA 2001, May 13–17, 2001, pp. 5-12.

[8] R. Chan, T. Chan and C. Wong, Cosine transform based preconditioners for total variation deblurring, IEEE Trans. Image Proc., 8 (1999), pp. 1472-1478.

[9] R. Chan, T. Chan, L. Shen and Z. Shen, Wavelet algorithms for high-resolution image reconstruction, SIAM Journal on Scientific Computing, 24 (2003) pp. 1408-1432.

[10] R. Chan, S. Riemenschneider, L. Shen and Z. Shen, Tight frame: An efficient way for high-resolution image reconstruction, Applied and Computational Harmonic Analysis, 17 (2004), pp. 91-115.

[11] T. Chan, S. Esedoglu, F. Park, and A. Yip, Recent developments in total variation image restoration, in N. Paragios, Y. Chen, and O. Faugeras, editors, Handbook of Mathematical Models in Computer Vision, pp. 17-32. Springer, 2005.

[12] T. Chan and J. Shen, Mathematical models for local non-texture inpainting, SIAM Journal on Applied Mathematics, 62 (2001), pp. 1019-1043.

[13] T. Chan, A. Yip, and F. Park, Simultaneous image inpainting and blind deconvolution, International Journal of Imaging Systems and Technology, 15:1 (2005), pp. 92-102.

[14] S. Chaudhuri, Editor, Super-Resolution Imaging, Kluwer Academic Publisher, Boston, 2001.

[15] M. Elad and A. Feuer, Super-resolution restoration of an image sequence, IEEE Transaction on Image Processing, 8 (1999), pp. 387-395.

[16] M. Elad and Y. Hel-Or, A fast superresolution reconstruction algorithm for pure translational motion and common space-invariant blur, IEEE Transaction on Image Processing, 10 (2001), pp. 1187-1193.

[17] G. Golub, M. Heath, and G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics, 21 (1979), pp. 215V223.

[18] G. Golub and C. Van Loan, Matrix Computations, 3rd ed., The Johns Hopkins University Press, 1996.

[19] P. C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, SIAM Rev., 34 (1992), pp. 561V580.

[20] R. Hardie, K. Barnard, J. Bognar, E. Armstrong, and Edward A. Watson, High-resolution image reconstruction from a sequence of rotated and rranslated frames and its application to an infrared imaging system, Optical Engineering, 37 (1998), pp. 247-260.

[21] M. Irani and S. Peleg, Improving resolution by image registration, CVGIP: Graphical Models and Image Processing, 53 (1991), pp. 231-239.

[22] S. Kim, N. Bose and H. Valenzuela, Recursive reconstruction of high resolution image from noisy undersampled multiframes, IEEE Trans. on Acoust., Speech, and Signal Process., 38 (1990), pp. 1013-1027.

[23] S. Kim and N. Bose, Reconstruction of 2-D bandlimited discrete signals from nonuniform samples, IEE Proceedings, 137, 3, Part F, (1990), pp. 197-203.

[24] S. Kim and W. Su, Recursive high-resolution reconstruction of blurred multiframe Images, IEEE Trans. on Image Proc., 2 (1993), pp. 534-539.

[25] P. Koeck, Ins and outs of digital electron microscopy, Microscopy Research and Technique, 49, (2000), pp. 217-223.

[26] L. Kolotilina and A. Yeremin, Factorized sparse approximate inverse preconditionings i. theory, SIAM J. Matrix Anal. and Appl., 14 (1993), pp. 45-58.

[27] S. Lertrattanapanich and N. Bose, Latest results on high-resolution reconstruction from video sequences, Technical Report of IEICE, DSP99-140, The Institution of Electronic, Information and Communication Engineers, Japan, December 1999, pp. 59-65.

[28] S. Lertrattanapanich and N. Bose, High resolution image formation from low resolution frames using Delaunay triangulation, IEEE Transactions on Image Processing, 11 (2002), pp. 1427-1441.

[29] F. Lin, M. Ng, and W. Ching. Factorized banded inverse preconditioners for matrices with Toeplitz structure, SIAM J. Sci. Comput., 26 (2005), pp. 1852-1870.

[30] S. Mann and R. Picard, Video orbits of the projective group: a simple approach to featureless estimation of parameters, IEEE Transactions on Image Processing, 6 (1997), pp. 1281-1295.

[31] S. Masnou and J. Morel, Level-lines based disocclusion, in Proc. 5th IEEE Int. Conf. on Image Process., pp. 259-263, Chicago, 1998.

[32] M. Ng and N. Bose, Analysis of Displacement errors in high-resolution image reconstruction with multisensors, IEEE Trans. on Circuits and Systems 1: Fundamental Theory and Applications, 49 (2002), pp. 806-813.

[33] M. Ng and N. Bose, Mathematical analysis of super-Resolution methodology, IEEE Signal Processing Magazine, 20 (2003), pp. 49-61.

[34] M. Ng, N. Bose and R. Chan, Guest editorial of high-resolution image reconstruction, International Journal of Imaging System and Technology, 14 (2004), pp. 35-89.

[35] M. Ng, N. Bose and J. Koo, Constrained total least squares computations for high resolution image reconstruction with multisensors, International Journal of Imaging Systems and Technology, 12 (2002), pp. 35-42.

[36] M. Ng, T. Chan, M. Kang and P. Milanfar, Guest editorial of Special Issue on Super-resolution imaging: analysis, algorithms, and applications, EURASIP Journal on Applied Signal Processing, 2006, to appear.

[37] M. Ng, R. Chan, T. Chan and A. Yip, Cosine transform preconditioners for high resolution image reconstruction, Linear Algebra and its Applications, 316 (2000), pp. 89-104.

[38] M. Ng, E. Lam and C. Tong, Guest editorial of Special Issue on the International Conference on Superresolution imaging, Multidimensional Systems and Signal Processing, 2006, to appear.

[39] M. Ng, C. Sze and S. Yung, Wavelet Algorithms for Deblurring Models, International Journal of Imaging Systems and Technology, 14 (2004), pp. 113-121.

[40] M. Ng and K. N. Sze, Preconditioned iterative methods for super-resolution image reconstruction with multisensors, in Symposium on Advanced Signal Processing: Algorithms, Architectures and Implementations, Franklin Luk, ed., Proceedings to the SPIE, 4116 (2000), pp. 396-405, San Diego CA, July, 2000.

[41] M. Ng and A. Yau. Super-resolution image restoration from blurred low-resolution image, Journal of Mathematical Imaging and Vision, 23 (2005), pp. 367-378.

[42] M. Ng and A. Yip, A fast MAP algorithm for high-resolution image reconstruction with multisensors, Multidimensional Systems and Signal Processing, 12 (2001), pp. 143-164.

[43] N. Nguyen and M. Peyman. A wavelet-based interpolation-restoration method for super-resolution (wavelet superresolution), Circuits, Systems, and Signal Processing, 19 (2000), pp. 321-338.

[44] S. Park, M. Park and M. Kang, Super-resolution image reconstruction: a technicial overview, IEEE Signal Processing Magazine, Special Issue of Superresolution Image Reconstruction, 20 (2003), pp. 21-36.

[45] S. Peled and Y. Yeshurun, Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging, Magnetic Resonance in Medicine, 45 (2001), pp. 29-35.

[46] L. Poletto and P. Nicolosi, Enhancing the spatial resolution of a two-dimensional discrete array detector, Engineering, 38 (1999), pp. 1748-1757.

[47] S. Rhee and M. Kang, Discrete cosine transform based regularized high-resolution image reconstruction algorithm, Optical Engineering, 38(8), August (1999), pp. 1348-1356.

[48] D. Rajan and S. Chaudhuri, An MRF-based approach to generation of super-resoltuion images from blurred observations, Journal of Mathematical Imaging and Vision, 16 (2002), pp. 5-15.

[49] L. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D, 60 (1992), pp. 259-268.

[50] K. Sauer and J. Allebach, Iterative reconstruction of band-limited images from nonuniformly spaced samples, IEEE Trans. on Circuits & Systems, Vol. 34 (1987), pp. 1497-1506.

[51] C. Srinivas and N. Srinath, A stochastic model based approach for simultaneous restoration of multiple mis-registered images, Proc. SPIE, 1360 (1990), pp. 1348-1356.

[52] H. Stark and P. Oskoui, High-resolution image recovery from image-plane arrays using convex-projections, J. Opt. Soc. Amer. A, 6 (1989), pp. 1715-1726.

[53] David Strong and T. Chan, Spatially and Scale Adaptive Total Variation Based Regularization and Anisotropic Diffusion in Image Processing, UCLA Math Department CAM Report, 96-46.

[54] W. Tang. Toward an effective sparse approximate inverse preconditioner, SIAM J. Matrix Anal. and Appl., 20 (1998), pp. 970-986.

[55] A. Tekalp, M. Ozkan and M. Sezan, High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration, in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, San Francisco, USA, 1992, pp. 169-172.

[56] A. Tekalp, M. Ozkan and M. Sezan, Superresolution video reconstruction with arbitrary sampling lattices and non-zero aperture time, IEEE Trans. Image Proc., 6, (1997), pp. 1064-1076,

[57] R. Tsai and T. Huang, Multiframe Image Restoration and Registration, Advances in Computer Vision and Image Processing: Image Reconstruction from Incomplete Observations, Thomas S. Huang, Ed., London, 1984, vol. 1, pp. 317-339, JAI Press.

[58] H. Ur and D. Gross, Improved resolution from subpixel shifted pictures, CVGIP: Graphical Models and Image Processing, Vol. 54, (1992), pp. 181-186.

[59] C. Vogel and M. Oman, Iterative methods for total variation denoising, SIAM J. Sci. Statist. Comput., 17 (1996), pp. 227-238.

[60] H. Voss and U. Eckhardt, Linear convergence of generalized Weiszfeld's method, Computing, 25 (1980), pp. 243-251.

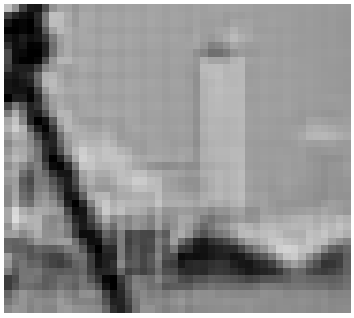
(a)

(b)

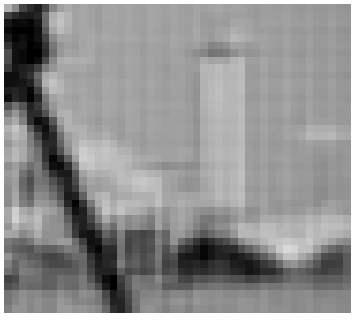
(c)

Figure 11: Comparison of regularization functionals. Images in left column are restored by TV regularization. Images in right column are restored by H^1 regularization. 11(a) Multi-frame without blur. 11(b) Multi-frame with single blur. 11(c) Multi-frame with multi-blur.

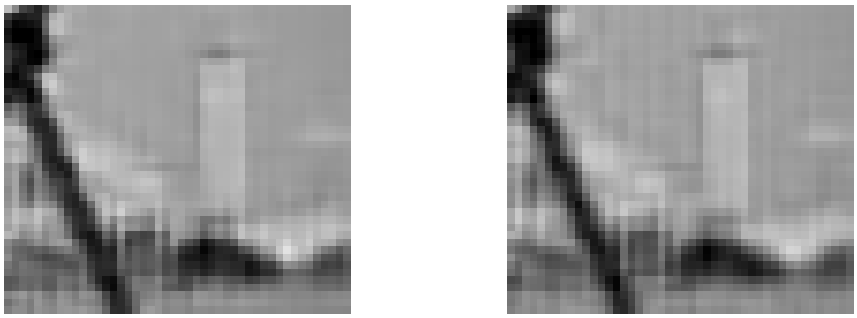
(a)



(b)



(c)



(d)

Figure 12: Detailed images of from two regularization functionals. Images in left column are restored by TV regularization. Images in right column are restored by H^1 regularization. 12(a) Original image. 12(b) Multi-frame without blur. 12(c) Multi-frame with single blur. 12(d) Multi-frame with multi-blur.