# A REGULARIZATION FRAMEWORK FOR JOINT BLUR ESTIMATION AND SUPER-RESOLUTION OF VIDEO SEQUENCES

*Hu He, Lisimachos P. Kondi*

State University of New York at Buffalo
Dept. of Electrical Engineering
Buffalo, NY 14260

### ABSTRACT

In traditional digital image restoration, the blurring process of the optic is assumed known. Many previous research efforts have been trying to reconstruct the degraded image or video sequence with either partially known or totally unknown point spread function (PSF) of the optical lens, which is commonly called the blind deconvolution problem. Many methods have been proposed in the application to image restoration. However, little work has been done in the super-resolution scenario. In this paper, we propose a generalized framework of regularized image/video iterative blind deconvolution / super-resolution (IBD-SR) algorithm, using some information from the more matured blind deconvolution techniques from image restoration. The initial estimates for the image restoration help the iterative image/video super-resolution algorithm converge faster and be stable. Experimental results are presented and conclusions are drawn.

## 1. INTRODUCTION

The goal of super-resolution is to estimate a high-resolution image from a sequence of low-resolution images while also compensating for blurring due to the point spread function of the camera lens and the effect of the finite size of the photo-detectors, as well as additive noise introduced by the capturing process. Super-resolution using multiple frames is possible when there exists subpixel motion between the captured frames. Thus, each of the frames provides a unique look into the scene. The problem of super-resolution is an active research area [1], [2], [3], [4], [5], [6], [7]. In traditional digital image restoration and super-resolution, the blurring process of the optic is assumed known. Many recent research efforts have been trying to reconstruct the degraded image or video sequence with either partially known or totally unknown point spread function (PSF) of the optical lens, which is commonly called the blind deconvolution problem. A lot of research has been performed on the blind deconvolution problem for image restoration. In our review of previous work, we focus on the regularized blind deconvolution methods [8], [9]. There are many advantages of this class of methods, such as, the fact that no parametric models for either the image or the PSF are required, the low computation complexity, and the robustness to noise. In [8], the piecewise smoothness of both the image and the PSF were adapted in the regularized cost function. The introduction of these constraints directly in the iterative procedure provides effective blind restoration quality. An estimation of PSF support was also proposed using thresholding and pruning. However, this results can not be directly applied to super-resolution because the additional degradations may introduce instability to the algorithm. A similar method was proposed in [9], with Total Variance (TV) norm replacing the H norm, but the method is limited to uniform blurs that have sharp edges.

In this paper, we propose a generalized framework of the regularized iterative blind deconvolution / super-resolution (IBD-SR) algorithm, using some information from the more matured blind deconvolution techniques from image restoration. The PSF blur is jointly estimated along with the update of the high-resolution image. The initial estimates for blind image restoration make the algorithm start from a good convergence region. Our algorithm is shown to be both computational efficient and stable. The rest of the paper is organized as follows. In section 2, the regularized super-resolution algorithm is presented with unknown but piecewise smooth PSF taken into consideration. The cost function now has two regularization parameters, one for the high-resolution (HR) image and one for the PSF. In section 3, the information from the blind deconvolution techniques in image restoration is used to determine the initial estimates of image/video super-resolution. In section 4, experimental results are presented and finally conclusions are drawn in section 5.

## 2. REGULARIZED SIMULTANEOUS BLIND DECONVOLUTION / SUPER-RESOLUTION TECHNIQUE

### 2.1. Observation Model

The image degradation process is modeled by motion, linear blur, subsampling by pixel averaging and an additive Gaussian noise process [1], [2], [3], [4], [5]. All vectors are ordered lexicographically. Assume that $p$ low-resolution frames are observed, each of size $N_1 \times N_2$. The desired high-resolution image $\mathbf{z}$ is of size $N = L_1N_1 \times L_2N_2$ where $L_1$ and $L_2$ represent the down-sampling factors in the horizontal and vertical directions, respectively. Thus, the observed low-resolution images are related to the high resolution image through motion shift, blurring and subsampling. Let the $k$th low-resolution frame be denoted as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,M}]^T$ for $k = 1, 2, \ldots, p$ where $M = N_1N_2$. The full set of $p$ observed low-resolution images can be denoted as

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_p^T]^T = [y_1, y_2, \ldots, y_{pM}]^T. \quad (1)$$

The observed low resolution frames are related to the high-resolution image through the following model:

$$y_{k,m} = \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) z_r + \eta_{k,m}, \quad (2)$$

0-7803-9134-9/05/$20.00 ©2005 IEEE

for $m = 1, 2, \ldots, M$ and $k = 1, 2, \ldots, p$. The weight $w_{k,m,r}(\mathbf{s}_k)$ represents the "contribution" of the $r$th high-resolution pixel to the $m$th low resolution observed pixel of the $k$th frame. The vector $\mathbf{s}_k = [s_{k,1}, s_{k,2}, \ldots, s_{k,K}]^T$, is the $K$ registration parameters for frame $k$, measured in reference to a fixed high resolution grid. The term $\eta_{k,m}$ represents additive noise samples that are assumed to be independent and identically distributed (i.i.d.) Gaussian noise samples with variance $\sigma_\eta^2$. The system can be modeled in matrix notation as

$$\mathbf{y} = \mathbf{W_z z} + \mathbf{n}. \quad (3)$$

In equation (3), the degradation matrix

$$\mathbf{W_z} = [\mathbf{W_{z,1}}, \mathbf{W_{z,2}}, \cdots, \mathbf{W}_{\mathbf{z},p}]^T \quad (4)$$

performs the operation of motion, blur and subsampling. Therefore $\mathbf{W_z}$ for frame $k$ can be written as:

$$\mathbf{W}_{\mathbf{z},k} = \mathbf{S}\mathbf{B}_k\mathbf{M}_k, \quad (5)$$

where $\mathbf{S}$ is the $N_1N_2 \times N$ subsampling matrix, $\mathbf{B}_k$ is the $N \times N$ blurring matrix, and $\mathbf{M}_k$ is the motion matrix. The PSF of blurring is assumed to be space-invariant, normalized and having non-negative elements within a 2-D rectangular support $l_1 \times l_2$ . The lexicographically ordered PSF $\mathbf{b}_k = [b_{k,1}, b_{k,2}, \cdots, b_{k,l_1l_2}]^T$, $k = 1, \cdots, p$, has the following properties:

$$\begin{cases} b_{k,u} \geq 0, u = 1, \cdots, l_1l_2 \\ ||\mathbf{b}_k|| = \sum_{r=1}^{l_1l_2} b_{k,u} = 1 \end{cases} \quad (6)$$

In this paper, we also assume global translation shift of the motion. This approximation will lose some generality, but is widely applicable for real data. Under these assumptions, $\mathbf{W}_{\mathbf{z},k} = \mathbf{S}\mathbf{M}_k\mathbf{B}_k$ and $\mathbf{B}_k$ is a Toeplitz matrix formed by $\mathbf{b}_k$. Thus, each frame can be modeled as

$$\mathbf{y}_k = \mathbf{W}_{\mathbf{z},k}\mathbf{z} + \mathbf{n}_k = \mathbf{S}\mathbf{M}_k\mathbf{B}_k\mathbf{z} + \mathbf{n}_k. \quad (7)$$

Also, $\mathbf{B}_k\mathbf{z}$ , the 2-D convolution of the PSF and the HR image, can be written as $\mathbf{Z}\mathbf{b}_k$ , by introducing a $N \times l_1l_2$ matrix $\mathbf{Z}$, whose columns are formed by circular shifting of vector $\mathbf{z}^T$. If we further define $\mathbf{W}_{\mathbf{b},k} = \mathbf{S}\mathbf{M}_k\mathbf{Z}$, Equation (7) can also be expressed as

$$\mathbf{y}_k = \mathbf{S}\mathbf{M}_k\mathbf{Z}\mathbf{b}_k + \mathbf{n}_k = \mathbf{W}_{\mathbf{b},k}\mathbf{b}_k + \mathbf{n}_k. \quad (8)$$

Due to the existence of additive noise, the deconvolution / super-resolution problem is a typical ill-posed problem. To overcome this problem, regularization can be used. We have done some research in previous work in [3], [4], with the assumption that the PSF is well known. When the PSF is unknown, we should also take the piecewise smoothness property of the PSF into consideration, i.e., form the following regularized cost function:

$$L(\mathbf{z}, \mathbf{b}_k) = \sum_{m=1}^{p} \{||\mathbf{y}_k - \mathbf{W}_{\mathbf{z},k}\mathbf{z}||^2\} + \alpha_{1,k}||\mathbf{D}\mathbf{z}||^2 + \alpha_{2,k}||\mathbf{D}\mathbf{b}_k||^2. \quad (9)$$

or

$$L(\mathbf{z}, \mathbf{b}_k) = \sum_{m=1}^{p} \{||\mathbf{y}_k - \mathbf{W}_{\mathbf{b},k}\mathbf{b}_k||^2\} + \alpha_{1,k}||\mathbf{D}\mathbf{z}||^2 + \alpha_{2,k}||\mathbf{D}\mathbf{b}_k||^2. \quad (10)$$

to minimize, where $\mathbf{D}$ is a high-pass filter formed by the 2-D Laplacian kernel. The cost function in Equation (9) or (10) can be approximately minimized via minimizing the following two functions alternatively, i.e., in a cyclic coordinate-descent optimization procedure.

$$L_1(\mathbf{z}) = \sum_{m=1}^{p} \{||\mathbf{y}_k - \mathbf{W}_{\mathbf{z},k}\mathbf{z}||^2\} + \alpha_{1,k}||\mathbf{D}\mathbf{z}||^2, \quad (11)$$

and

$$L_2(\mathbf{b}_k) = \sum_{m=1}^{p} \{||\mathbf{y}_k - \mathbf{W}_{\mathbf{b},k}\mathbf{b}_k||^2\} + \alpha_{2,k}||\mathbf{D}\mathbf{b}_k||^2. \quad (12)$$

It is difficult to get the minimizer of the high-resolution image $\mathbf{z}$ directly. To minimize the cost function in (11), an iterative algorithm can be used:

$$\hat{z}_r^{n+1} = \hat{z}_r^n - \epsilon g_r(\hat{\mathbf{z}}_n), r = 1, \cdots, N. \quad (13)$$

where the gradient $g_r$ is

$$g_r = \frac{\partial L_1(\mathbf{z})}{\partial z_r}. \quad (14)$$

In [4], we have shown that a good choice for the regularization parameter is:

$$\alpha_{1,k} = \frac{||\mathbf{y}_k - \mathbf{W}_{\mathbf{z},k}\mathbf{z}||^2}{2||\mathbf{y}_k||^2 - ||\mathbf{D}\mathbf{z}||^2}, \quad (15)$$

and

$$\epsilon = \frac{2}{p}\left(\frac{(L_1L_2)}{(L_1L_2)\phi_{max}(\mathbf{D}^T\mathbf{D}) + 1}\right), \quad (16)$$

where $\phi_{max}(\cdot)$ is the maximum eigenvalue of a matrix. Due to the relatively smaller size of the support size $l_1l_2$ and fewer unknown components in the PSF, we can solve $\mathbf{b}_k$ from the cost function (12) as:

$$\mathbf{b}_k = (\mathbf{W}_{\mathbf{b},k}^T\mathbf{W}_{\mathbf{b},k} + \alpha_{2,k}\mathbf{D}^T\mathbf{D})^{-1}(\mathbf{W}_{\mathbf{b},k}^T\mathbf{y}_k). \quad (17)$$

The regularization parameter $\alpha_2$ of the PSF is obtained via a trial and error method.

### 2.2. Determination of the Initial Estimates in Super-resolution using Some Information from Image Restoration

It is very important to choose the initial estimates of the proposed regularization super-resolution techniques. The alternative update of the high-resolution image and the PSF blur may introduce instability to the algorithm. It is necessary to run the algorithm in the region near the optimal solutions. Therefore, we use the more matured techniques of image restoration, and propose the following procedure to obtain the good initial estimates of both the high-resolution image and the PSF blur.

First, we reuse the cost function (11) in an image restoration case:

$$L(\mathbf{z}_k) = ||I_\uparrow(\mathbf{y}_k) - \mathbf{B}_k\mathbf{z}_k||^2 + \alpha_k||\mathbf{D}\mathbf{z}_k||^2, \quad (18)$$

where $I_\uparrow(\cdot)$ is an interpolation function, such as bilinear or bicubic interpolation function. The $\mathbf{z}_k$ is the motion-shifted version of the high-resolution $\mathbf{z}$ for channel (low-resolution image) $k$, i.e., $\mathbf{z}_k = \mathbf{M}_k\mathbf{z}$. This modification makes the motion estimation is not necessary in the minimization of cost function (18). Each channel can estimate its own estimation pair, $\mathbf{B}_k$ and $\mathbf{z}_k$. After interpolation, $I_\uparrow(\mathbf{y}_k)$ has the same size of $\mathbf{z}$, i.e., $N \times 1$. Compared to the cost function (11), we also drop the summation of $p$ frames and change the degradation matrix from $\mathbf{W}_{\mathbf{z},k}$ to $\mathbf{B}_k$ in (18). These changes are made based on: (1) only frame $k$ is needed in the cost function; (2) the subsampling operation $\mathbf{S}$ is dropped due to interpolation; (3) the subpixel motion $\mathbf{M}_k$ is included in $\mathbf{z}_k$, a motion-shifted version of $\mathbf{z}$.

Similar reusage of Equations (15)-(17) can be made for the image restoration.

These changes make the initial estimates of both the high-resolution image and the PSF blur much easier and fast. Any optimization technique can be used for cost function (18). The $k$ motion-shifted versions of the high-resolution image, $\mathbf{z}_k$, can be used to estimate the motion vectors between low-resolution frames. The initial estimate of high-resolution image $\mathbf{z}$, can be set to the high-resolution reconstruction from the reference frame (with motion vector $[0, 0]$), or fused from the $k$ motion-shifted solutions $\mathbf{z}_k$ after motion compensation. If some other prior information is available, the above procedure can be even faster. For example, if we know the PSF blur is Gaussian type with Gaussian variance $\sigma_b^2$ as the parameter, we can run the trial and error method of $\sigma_b^2$ for (11) and choose the solution with best visual reconstruction.

A better interpolation function $I_\uparrow(\cdot)$ from multiple frames may be helpful, but bilinear or bicubic interpolation of a single frame $k$ usually is enough to obtain fairly good initial estimates.

## 3. EXPERIMENTAL RESULTS

A number of experiments were conducted, some of which are presented here. To test the performance of our algorithms, we first use the $256 \times 256$ "Cameraman" test image for a synthetic test. The PSF is a Gaussian PSF with support size $15 \times 15$ and standard deviation $\sigma = 1.7$. Three cases, Case I-III, as listed in Table 1, are tested. The global shift $\mathbf{s}_k^T$ belongs to the vectors generated from the given Cartesian product listed in the table. The first low-resolution frame is selected as the reference frame. The initial estimates of the high-resolution image $\mathbf{z}$ and PSF blur are obtained using the techniques developed in previous section. The coordinate-descent method is then applied to get a more enhanced HR image $\mathbf{z}$ and a better estimated PSF blur $\mathbf{b}_k$. The alternate updating of the PSF and the HR image can continue until either PSF blur $\mathbf{b}_k$ or the HR image $\mathbf{z}$ converges.

The PSNR of the reconstructed images for "Cameraman" using three methods (Bilinear interpolation (BI), MAP based super-resolution with known PSF and Gaussian-Markov Random Field (GMRF) as the image prior [3], [4], and the proposed method) are listed in Table 2. Here, PSNR$_{HR}$ is $10\log_{10}\frac{255^2}{MSE_{HR}}$ and PSNR$_{PSF}$ is $10\log_{10}\frac{1}{MSE_{PSF}}$, where MSE stands for the mean squared error between the original HR / PSF and the estimated HR / PSF. The reconstructed "Cameraman" image from bilinear interpolation, GMRF with the PSF known [4], the proposed method in case II are shown in Figs. 2, 3, and 4, respectively. The original PSF and reconstructed PSF are listed in Figs. 5 and 6. From the results, we can see that our algorithm not only stabilizes the deconvolution procedure, but generates good reconstruction results.

## 4. CONCLUSION

We propose a regularized iterative image/video blind deconvolution algorithm using two regularization parameters, one for the HR


**Fig. 1**. Original Cameraman image.

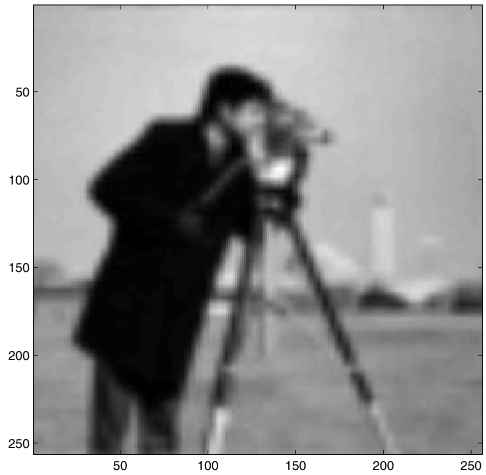


**Fig. 2**. Bilinear interpolation of the first low-resolution frame of Cameraman.

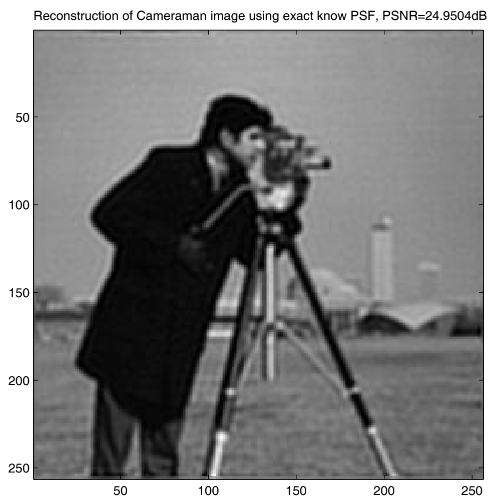


**Fig. 3**. Reconstruction of Cameraman image using exactly known PSF.

**Table 1**. Three cases of synthetic test for "Cameraman".

| | $L_1, L_2$ | $p$ | $\mathbf{s}_k^T = [s_{k,1}, s_{k,2}]$ | $\sigma_\eta^2$ |
|---|---|---|---|---|
| Case I | 1 | 1 | $\{0\} \times \{0\}$ | 1 |
| Case II | 2 | 4 | $\{0,1\} \times \{0,1\}$ | 1 |
| Case III | 4 | 16 | $\{0,1,2,3\} \times \{0,1,2,3\}$ | 1 |

**Table 2**. Results of "Cameraman" using the three methods.

| PSNR (dB) | BI | GMRF | Proposed: $\text{PSNR}_{HR}$ | Proposed: $\text{PSNR}_{PSF}$ |
|---|---|---|---|---|
| Case I | 23.1319 | 24.9287 | 24.7959 | 61.6270 |
| Case II | 22.5218 | 24.9504 | 24.6947 | 60.1632 |
| Case III | 21.2616 | 24.0632 | 23.9731 | 44.6216 |

image and one for the PSF. we also propose a generalized framework to use some information from the more matured blind deconvolution techniques from image restoration. The initial estimates from blind image restoration help the super-resolution algorithm converge faster and be stable.

## 5. REFERENCES


[1] R. R. Schultz and R. L. Stevenson, "A bayesian approach to image expansion for improved definition," *IEEE Trans. Image Processing*, vol. 3, pp. 233–242, May 1994.

[2] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Processing*, vol. 5, pp. 996–1011, June 1996.

[3] H. He and L. P. Kondi, "MAP based resolution enhancement of video sequences using a Huber-Markov random field image prior model," *IEEE International Conference on Image Processing*, vol. 2, pp. 933–936, 2003.

[4] H. He and L. P. Kondi, "Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter," *Journal of Electronic Imaging*, vol. 13, pp. 586–596, July 2004.

[5] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Processing*, vol. 6, pp. 1621–1633, Dec. 1997.

[6] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. Image Processing*, vol. 10, pp. 1299–1308, Sept. 2001.

[7] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Trans. Image Processing*, vol. 10, pp. 1187–1193, Aug. 2001.

[8] Y. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," *IEEE Trans. Image Processing*, vol. 5, pp. 416–428, March 1996.

[9] T. F. Chan and Chiu-Kwong Wong, "Total variation blind deconvolution," *IEEE Trans. Image Processing*, vol. 7, pp. 370–375, March 1998.



**Fig. 4**. Reconstruction of Cameraman image using unknown PSF.

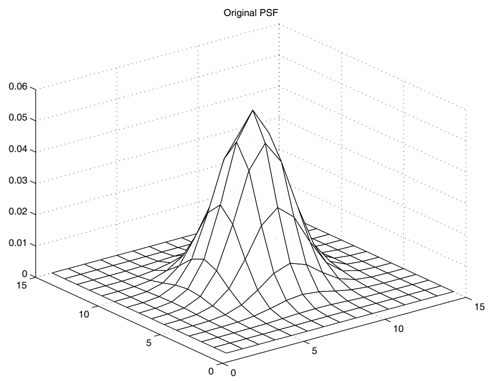


**Fig. 5**. Original PSF.

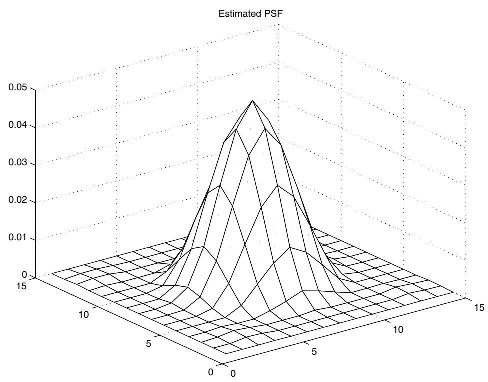


**Fig. 6**. Reconstructed PSF.