

Military Technical College
Kobry El-Kobbah,
Cairo, Egypt



7th International Conference
on Electrical Engineering
ICEENG 2010

Learning-Based Image Super-Resolution with Directional Total Variation

By

Osama. A. Omer*

Abstract:

We propose a super-resolution algorithm based on local adaptation. In the proposed algorithm, the mapping function from the low-resolution images to high-resolution image is estimated by adaptation. Moreover, the property of the high-resolution image is learned and incorporated in a regularization-based restoration. The proposed regularization function is used as a general directional total variation with adaptive weights. The adaptive weights of the directional total variation are estimated based on the property of the partially reconstructed high-resolution image. The regularization function can be thought as a linear combination of smoothness in different directions. The convexity conditions as well as the convergence conditions are studied for the proposed algorithm.

Keywords:

Super-resolution, image fusion, restoration, directional total variation, regularization

* Department of Electrical Engineering, South Valley University, Aswan
Email: usama.abdelaziz@asw-eng.edu.eg

1. Introduction:

The problem of image super-resolution (SR), where one obtains high-resolution (HR) image(s) from a set of low-resolution (LR) images, has many applications in image processing and computer vision fields, including remote sensing and medical images. The task of resolution enhancement is typically to increase spatial resolution which can assist the recognition in LR images.

Methods of image SR can be divided into two classes. Namely, multi-images SR and single image SR. In the former class, the HR image is obtained by fusing information from multiple LR images to get a single image with higher resolution while in the later class, one extracts HR image details from a single LR image, which cannot be achieved by simple interpolation and sharpening.

Many approaches for image SR from multiple LR images have been presented in the last two decades by applying different ideas (for review see [1]) including recursive least squares (RLS), iterative back-projection (IBP), maximum likelihood (ML), maximum a posteriori (MAP), projection onto convex sets (POCS), non-uniform interpolation, FIR filter, machine learning and neural networks. Among these algorithms, the two steps-based SR algorithms [2, 3] are reviewed and extended in this paper for its simplicity.

In [2, 3], the SR process has been performed by two steps, namely, the fusion step and restoration step. In [2], the LR images are fused by shift-and-add, where the up-sampled and warped images are averaged to get a blurred version of the HR image which is restored in the restoration step. In this algorithm, it is assumed that all the LR images contain HR information. In fact this assumption is not always correct due to the registration error. On the other hand, the median-shift-and-add (MSA) is used in [3] to fuse the LR images. Indeed, taking median of the up-sampled and warped images is more robust against the registration error; however, median is effective in case that less than 50 % are error pixels which are not guaranteed in case of two sources of noise (the registration error and additive noise).

In addition, the HR image is reconstructed by performing a restoration step to the fused images. Because of the ill-posedness of the restoration problem, typically the ill-posed problem is solved by adding constraint(s) to the solution by adding a regularization term to the cost function. In [2], the Tikhonov regularization is used as a regularization function. In [3], the bilateral total variation (BTV) is used as a regularization function. The use of Tikhonov regularization implicitly assumes smoothness of the HR image in all direction which is not always correct. On the other

hand, BTV regularization can overcome smoothness cross sharp edges by adjusting the bilateral filters' parameters. However, the main problem of BTV is the great dependency on the bilateral filters' parameters. Moreover, the weight for smoothness in each direction has a fixed value for constant bilateral parameters regardless to the image's property.

Hence, the motivation of this work is to develop an unsupervised learning-based SR algorithm that can overcome the problems of both fusion step and restoration step. To do that, the mapping function between the LR and HR images is estimated by learning in the fusion step which can overcome the problem of using fixed mapping function. Moreover, in the restoration step, rather than using Tikhonov or BTV as a regularization function, a general *directional total variation* (DTV) function with adaptive weights is newly introduced in a regularization-based restoration. The weights for different direction are adaptively estimated based on the partially estimated HR image. This weight adaptation can sense the image's local properties and then overcomes the problem of using fixed smoothness weight for different direction.

The remaining sections of this paper are organized as follows. Section 2 describes the mathematical model for the super-resolution problem. Section 3 briefly reviews a state-of-the-art two steps-based SR algorithm and the problems of this algorithm are highlighted. The proposed learning-based fusion algorithm is introduced in Section 4. The directional-regularization-based restoration algorithm is presented in section 5. The proposed algorithm is generalized to batch-based SR in Section 6. Section 7 presents the simulation results on the fusion, restoration and SR of document image sequences. Finally, the conclusions are presented in Section 8.

2. Problem Description

Assume that K LR images of the same scene in Lexicographical order denoted by Y_k ($1 \leq k \leq K$), each containing M^2 pixels, are observed, and they are generated from the HR image denoted by X , containing L^2 pixels, where $L < M$. The observation of K LR images are modeled by the following degradation process:

$$Y_k = \mathbf{D}\mathbf{H}\mathbf{F}_k X + V_k \quad (1)$$

Where \mathbf{F}_k , \mathbf{H} and \mathbf{D} are the motion operator of the k^{th} image, the blurring operator (due to camera), and the down-sampling operator respectively, X is the unknown HR image, Y_k is the k^{th} observed LR image, and V_k is an additive random noise for the k^{th} image. The dimensions of \mathbf{F}_k , \mathbf{H} , \mathbf{D} , X , Y_k , V_k are $(L^2 \times L^2)$, $(L^2 \times L^2)$, $(M^2 \times L^2)$, $(L^2 \times 1)$, $(M^2 \times 1)$ and $(M^2 \times 1)$, respectively. Note; the capital bold-faced font

is used to describe matrices.

Throughout the paper, we assume that \mathbf{D} and \mathbf{H} are known and the additive noise is Gaussian with zero mean. Therefore the problem here in this paper is to find the original image X .

3. Two Steps-Based Super-Resolution

Traditionally, the SR problem is solved by using regularization-based algorithm [3] where BTV is used as the regularization term. The traditional cost function for SR is described as

$$J(X) = \lambda \sum_{l=-d}^d \sum_{m=-d}^d \alpha^{|l|+|m|} \|X - \mathbf{S}_x^l \mathbf{S}_y^m X\|_1 + \sum_{k=1}^K \|\mathbf{DHF}_k X - Y_k\|_1 \quad (2)$$

Where α is the bilateral filter's parameter which is always pre-determined manually and λ is the regularization parameter.

In the document images, motion can be approximated as pure translation or even it can be adjusted in the system to be pure translation. Then both blur operator and motion operator can commute (i.e. $\mathbf{HF}_k = \mathbf{F}_k \mathbf{H}$). Therefore, the cost function, described in (2), can be modified as

$$J(X) = \lambda \sum_{l=-d}^d \sum_{m=-d}^d \alpha^{|l|+|m|} \|X - \mathbf{S}_x^l \mathbf{S}_y^m X\|_1 + \sum_{k=1}^K \|\mathbf{DF}_k \mathbf{H}X - Y_k\|_1 \quad (3)$$

Assume that $Z = \mathbf{H}X$ is the blurred version of the HR image. The optimization problem described in (3) can be separated into two sub-problems, that is, fusion and restoration:

$$J_1(Z) = \sum_{k=1}^K \|\mathbf{DF}_k Z - Y_k\|_1 \quad (4)$$

And

$$J_2(X) = \lambda \sum_{l=-d}^d \sum_{m=-d}^d \alpha^{|l|+|m|} \|X - \mathbf{S}_x^l \mathbf{S}_y^m X\|_1 + \|\mathbf{A}(\mathbf{H}X - Z)\|_1 \quad (5)$$

Where \mathbf{A} is a diagonal matrix that contains the square root of the number of LR pixels corresponding each pixel position, \mathbf{S}_x^l and \mathbf{S}_y^m are the shifting operators in x and y by l and m respectively, and $((2d + 1) \times (2d + 1))$ is the support size on which

the smoothness is done. The optimum solution (Z) that minimizes (4) is found to be the median of the up-sampled and warped LR images [3], i.e.

$$Z(i) = \underset{k}{\text{median}}(F_k^T D^T Y_k)(i) \quad (6)$$

This solution implies a fixed mapping function between the LR images and blurred version of the HR image. Also, the reconstructed HR image by using (5) is highly depending on the choice of α and d .

4. Learning-Based Fusion

In this step, the LR images are fused to get a blurred version of the HR image, where the mapping function between the LR images and blurred HR image is learned. Assume that the required resolution enhancement factor is r , then, r^2 sub-grids need to be filled in the HR grid. Because of the assumption of pure translational motion, all pixel positions on each sub-grid have the same number of LR pixels to be fused. Therefore, it is reasonable to update the weights of each sub-grid independently. The mapping function is defined by the weighting parameters W as follows

$$\hat{Z}(i) = \sum_{k=1}^{\mathcal{K}_j} W_j(k) (F_k^T D^T Y_k)(i), \quad i \in G_j \quad (7)$$

Subject to

$$\sum_{k=1}^{\mathcal{K}_j} W_j(k) = 1$$

Where W_j is a $(\mathcal{K}_j \times 1)$ vector that weights the contribution of the LR pixels that

aligned to the sub-grid G_j and \mathcal{K}_j is the number of LR images aligned to that sub-grid.

The mapping parameters are learned so as to minimize the L_1 -norm of the difference between the observed LR images and the simulated LR images, where

the simulated LR images are generated using the observation model in (1) as follows:

$$W_j = \arg \min_{W_j} \sum_{k=1}^{\kappa_j} \left\| \mathbf{D}\mathbf{F}_k \hat{Z} - Y_k \right\|_1 \quad (8)$$

Therefore, the updating rule for parameters W_j are defined as:

$$W_j(n+1) = W_j(n) - \mu_{W_j} Y_i(n) \sum_{k=1}^{\kappa_j} \text{sign} \left((\mathbf{D}\mathbf{F}_k \hat{Z})(n) - Y_k(n) \right) \quad (9)$$

Where n is the sample's index of the sub-grid. The mapping parameters are updated at each pixel position of the HR grid. The updating process of the mapping parameters is described as follows; starting at position (0,0) of the sub-grid G_j , the mapping parameters are initialized as

$$W_j^0 = \left[\frac{1}{\kappa_j} \Lambda \quad \frac{1}{\kappa_j} \right]$$

Then the parameters are updated progressively according to a sequential scanning pattern for each sub-grid.

Adaptive Step Size

For its stability and fast convergence, the adaptive step size of the generalized NLMS algorithm [4] for L_1 -norm case is used to choose the step size of learning parameters. The step size of the mapping function's parameters is chosen as

$$\mu_{W_j} = \frac{\gamma}{c + \|R_{W_j}\|_1} \quad (10)$$

Where $c > 0$ and $R_{W_j} = \left[(\mathbf{F}_1^T \mathbf{D}^T Y_1)(n) \Lambda \quad (\mathbf{F}_{\kappa_j}^T \mathbf{D}^T Y_{\kappa_j})(n) \right]^T$.

5. Directional Regularization (DR) and Adaptation

Instead of assuming smoothness in all directions, by using Tikhonov regularization, or using constant values for λ and μ as in (5), we propose to use a general regularization function consisting of a linear combination of the smoothness in different directions with adaptive weight $\lambda_{l,m}(X)$. To do that, we propose to replace $\lambda^{||l+m||}$ by a generalized parameter, $\lambda_{l,m}(X) \geq 0$, which we call the **DR parameter** in direction (l, m) . We want to emphasize that $\lambda_{l,m}$ is not only a generalization of the conventional non-directional parameter, but also is a function of X . We will discuss how to define this function and how to adapt this parameter later. In fact this regularization function can be viewed as smoothness in different direction with different regularization parameter.

Adaptive DR-Based Restoration

The cost function of the proposed restoration algorithm with DR-based regularization is described as

$$J_3(\Gamma(X), X) = \sum_{l=-d}^d \sum_{m=-d}^d \Gamma_{l,m}(X) \|\mathbf{C}_{l,m} X\|_1 + \|\mathbf{H}X - Z\|_1 \quad (11)$$

Where $\|\cdot\|_1$ is the L₁-norm, $\Gamma_{l,m}(X)$ is the adaptive weight for direction (l,m) to be learned based on the partially restored image and $\mathbf{C}_{l,m}$ is a high pass operator that works in the direction (l,m).

Iterative Minimization

The cost function described in (11) is optimized by using steepest decent for restored image estimation and iterative approximation for DR parameters computation. The coefficients of the DR function ($\Gamma_{l,m}(X)$) are computed for each direction so as to preserve the convexity of the function $J_3(\Gamma(X), X)$. Assuming $\Gamma_{l,m}(X)$ as a linear function of $J_{l,m}(\Gamma_{l,m}(X), X)$ as

$$\Gamma_{l,m}(X) = \xi_{l,m} J_{l,m}(\Gamma_{l,m}, X) \quad (12)$$

Where

$$J_{l,m}(\Gamma_{l,m}, X) = \|\mathbf{H}X - Z\|_1 + \Gamma_{l,m} \|\mathbf{C}_{l,m} X\|_1 \quad (13)$$

Then

$$\Gamma_{l,m}(X) = \frac{\|\mathbf{H}X - Z\|_1}{\frac{1}{\xi_{l,m}} - \|\mathbf{C}_{l,m} X\|_1} \quad (14)$$

Therefore, the restored image is updated as

$$X^{(n+1)} = X^{(n)} - \beta \nabla_X J_3(\Gamma(X^{(n)}), X^{(n)}) \quad (15)$$

Where $\nabla_X J_3(\Gamma(X), X)$ is computed as follows:

$$\nabla_X J_3(\Gamma(X), X) = \mathbf{H}^T S_1 + \sum_{l=-d}^d \sum_{m=-d}^d \Gamma_{l,m}(X) \mathbf{C}_{l,m}^T S_2 + \sum_{l=-d}^d \sum_{m=-d}^d \nabla_X \Gamma_{l,m}(X) \|\mathbf{C}_{l,m}^T X\|_1 \quad (16)$$

$$\nabla_X \Gamma_{l,m}(X) = \left(\frac{1}{\xi_{l,m}} - \|\mathbf{C}_{l,m}^T X\|_1 \right) \mathbf{H}^T S_1 + \|\mathbf{H}X - Z\|_1 \mathbf{C}_{l,m}^T S_2 \left(\frac{1}{\xi_{l,m}} - \|\mathbf{C}_{l,m}^T X\|_1 \right)^2 \quad (17)$$

Where $S_1 = \text{sign}(\mathbf{H}X - Z)$ and $S_2 = \text{sign}(\mathbf{C}_{l,m} X)$.

6. Patch-Based Local Property Learning

In the previous section, the proposed algorithm was described so that the properties of the restored image are learned globally. However, the properties of the restored image are always changing locally. So, it is more suitable to adaptively change the DR parameters locally based on the properties of the local positions. To do that, DR parameters, (X) , are computed locally from overlapped local patches. Therefore, the cost function ($J_3(X, X)$), defined in (11) is modified as

$$J_4(\Gamma(X), X) = \sum_{i=1}^N \left\{ \|P_i(\mathbf{H}X - Z)\|_1 + \sum_{l=-d}^d \sum_{m=-d}^d \Gamma_{l,m}(X) \|P_i(\mathbf{C}_{l,m}X)\|_1 \right\} \quad (18)$$

Where N is the number of patches, and P_i is operator that selects the patch i . As shown in [5], the sufficient condition for convexity is

$$\frac{1}{\xi_{l,m,i}} > \|P_i(\mathbf{C}_{l,m}X)\|_1 \quad (19)$$

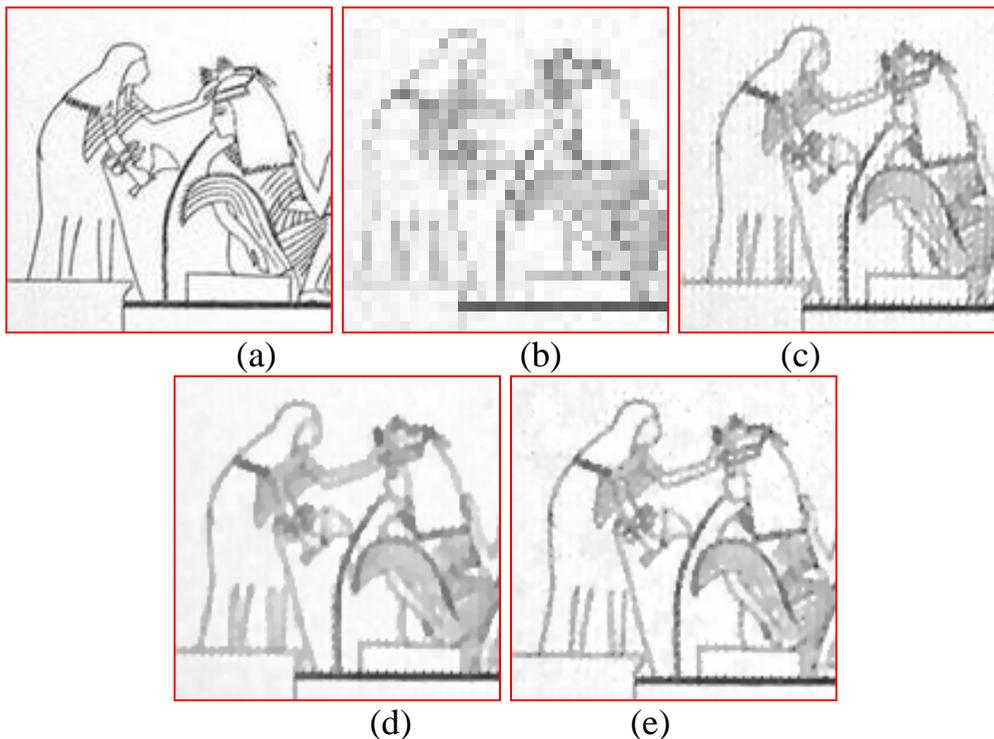


Figure (1): SR of Draw image sequence: (a) Original HR image, (b) LR image, Super-resolved image using; (c) MSA + BTV restoration [3] with worst selection of ξ , (d) MSA + BTV restoration [3] with best selection of ξ , (e) Learning-based fusion + patch-based

DR restoration algorithm

7. Simulation results

Experiment Setup

In the experiments, the LR images sequence of the document images are generated from the available HR images as follows. At first, the HR image is blurred with a Gaussian operator with kernel size equals 5×5 and variance equals 1. Then the blurred images are shifted with random values to generate K randomly shifted versions of the HR image ($K = 30$). Finally, the shifted versions of the HR images are down-sampled with factor equals 4 in each direction and an additive Gaussian noise with SNR equals 30 dB is added.

In all experiments, the blur operator is assumed to be known. The parameters used in the simulations are set as follows. Steepest decent is used for optimization. Motion is estimated by using Lucas-Kanade's Algorithm [7].

The maximum number of iterations are 100, $\alpha = 5$, $\beta = 0.5$, and the patch size is set as 11×11 with two overlapping pixels in each side. $\lambda_{l,m,i}$ is chosen to equal

$\frac{1}{2} \|P_i(Z)\|_1$, for all l,m values. In the evaluation, the reconstructed HR images using

learning-based fusion + DR-based restoration is compared with the best and worst results of MSA + BTV-based restoration algorithm. Also, for more fair comparison, in case of BTV-based restoration is adaptively estimated in a similar way as in [6].

SR Results

The overall SR algorithm including fusion and restoration steps is evaluated in this section. To test the efficiency of the proposed algorithm, it is compared with MSA-based fusion + BTV-based regularization [3].

In Figure 1, the results of the Draw sequence are shown. A zoomed part for the original HR and one of the LR images are shown in Figures 1a and 1b, respectively. The HR image by using MSA with BTV regularization for worst selection (corresponds to the highest MAE value) and best selection of are shown

in Figures 1c and 1d, respectively. Using MSA can be a good choice in case of less outliers contaminating LR images sequence. However, in case of heavy noisy sequences, MSA can fail to remove all noisy pixels as shown in some noisy pixels in Figure 1c. On the other hand, the proposed learning-based algorithm can learn the weights of the LR pixels to be included (or even completely excluded) in the HR image.

Moreover, a general prior property function is assumed in the regularization term, which is learned simultaneously to overcome the over-smoothness problem and remove contaminating noise, as shown in Figure 1e.

7. Conclusions:

In this paper, we presented a novel learning-based image SR algorithm. The proposed algorithm is divided into two main steps, the first is the non-iterative fusion step, and the second is the iterative restoration step. The fusion step is based on learning the mapping function from the LR to HR images. The restoration step is based on adapting the regularization parameter in each direction based on the partially reconstructed image. The convexity condition as well as the convergence conditions is introduced. The proposed algorithm is tested for document image super-resolution application.

References:

- [1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Magazine*, vol. 20, no. 3, pp. 21-36, May 2003.
- [2] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure transnational motion and common space invariant blur," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1187-1193, August 2001.
- [3] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Trans. Image Processing*, vol. 13, no.10, pp. 1327-1344, October 2004.
- [4] S. C. Douglas, "A family of normalized LMS algorithms," *IEEE Signal Process. Letters*, vol. 1, no. 3, pp. 49-51, March 1994.

- [5] O. A. Omer and T. Tanaka, "Image restoration based on adaptive directional regularization," *IEICE Trans. Fundamentals*, vol. E92-A, no.12, December 2009.
- [6] M. G. Kang and A. K. Katsaggelos, "General choice of the regularization functional in regularized image restoration," *IEEE Trans. Image Process.*, vol. 4, no. 5, pp. 594-602, May 1995.
- [7] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artificial Intelligence*, April 1981, pp. 674-679.