Int. J. Electron. Commun. (AEÜ) 70 (2016) 409–415

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Simultaneous edge preserving and noise mitigating image super-resolution algorithm

Gunnam Suryanarayana*, Ravindra Dhuli

School of Electronics Engineering, VIT University, Vellore 632014, Tamilnadu, India

ARTICLE INFO

Article history:
Received 21 February 2015
Accepted 28 December 2015

Keywords:
Diffusion based shock filter
Image super-resolution
Stationary wavelet transform (SWT)
Image denoising

ABSTRACT

State-of-the-art single image super-resolution (SISR) methods provide faithful reconstruction, but involve a training step using large database, demanding high computations. We propose a method which reduces the execution time significantly by eliminating the training process. To preserve the edges, stationary wavelet transform (SWT) is employed. Further image enhancement and noise sensitivity depletion is achieved using complex diffusion based shock filter by operating in the dual dominant mode. These filtered subbands are combined to generate a high resolution (HR) image. Further artifacts are removed by projecting onto a global image vector space iteratively. Experimental results show that the performance of the proposed method is superior to the existing methods.

1. Introduction

Super-resolution (SR) reconstruction is an inexpensive signal processing approach to gain a high resolution (HR) image from single or multiple degraded low resolution (LR) input images [1–21]. These algorithms overcome the drawbacks of low cost imaging sensors and meet the growing demand for HR displays. However, the problem of SR is under determined due to the unknown blurring operators. SR applications include Generic Image Enhancement, Medical Imaging, Satellite and Aerial Imaging, Infrared and Ultrasonic Imaging, Face Hallucination, Text Image Restoration, etc.

In conventional SR techniques [1,2], the HR image is obtained by fusing a number of LR frames aligned with some geometric shifts between them. However, these methods are inaccurate due to insufficient LR frames, ill-conditioned registration parameters and unstable motions between the LR images. On the other hand, single image based SR is the most challenging research domain. The classical methods for single image super-resolution (SISR) are mostly interpolation based [3–6,21]. Simple interpolation techniques such as nearest-neighbor, bilinear and bicubic produce images with blurred edges and undesired ringing and jagged artifacts. To overcome the inherent drawbacks in bicubic interpolation and recover the lost high frequency details, Xin and Michael [3] developed a relation between LR and HR images using the local covariance estimates of the input LR image. Lei and Xiaolin [4] defined two orthogonal sets of pixel estimates which are fused by the linear minimum mean square error estimation. In [6] Zhang and Wu improved the edge directed interpolations [3,4] using 2D piecewise auto regressive model. In [5] contourlet transform is employed to highlight the object boundaries in the HR image. However, these methods (except [6,21]) are not providing satisfactory results, in terms of various image quality measures [8].

Another category of SISR approach rely on machine learning algorithms [7–16]. They include a training strategy between several HR frames and their LR counter parts. Here SR algorithms need an apriori term acquired from the dictionary developed using the learning strategy. In [7] Fadili et al. developed expectation maximization algorithm which uses a sparsity prior for image inpainting and zooming. This mechanism yields better results for inpainting than image zooming. Sparsity is effectively used in many SR algorithms [7–16]. It develops the linear relationships between HR patches and their LR counter parts assuming two overcomplete dictionaries, one for the HR samples the other for their LR projections.

Weisheng Dong et al. [12,13,16] further improved the sparse coding techniques by introducing centralized sparse representation and non local autoregressive model. Li et al. [14] proposed a new model which uses a beta process prior for dictionary learning with some assumption of the sparsity invariance. But, it is restricted to the non-zero locations in the sparse representation vector. This issue is addressed in [15], which makes no invariance assumption by adapting a statistical prediction model. The sparse representation vector of HR patch is obtained from its LR version via minimum mean square error estimation. This approach has an effective

* Corresponding author. Tel.: +91 9985112223.
E-mail addresses: gunnam.suryanarayana@vit.ac.in (G. Suryanarayana), ravindradhuli@vit.ac.in (R. Dhuli).

http://dx.doi.org/10.1016/j.aeue.2015.12.020
1434-8411/

interpretation of feed forward neural network, which leads to better reconstruction quality. However, the above mentioned techniques need large training data sets which in turn increases the computational complexity.

A different class of SR approach exploits the wavelet transform. The wavelet techniques play a vital role in numerous signal and image processing applications. In the context of SR several algorithms have been proposed in the wavelet domain [8,17–19]. In the present work, we develop a wavelet based SR algorithm by employing shock filter. In this process we extract the edge information from input LR image. To overcome the inherent shift variance of the discrete wavelet transform (DWT) we choose stationary wavelet transform (SWT). The cycle spinning and undecimated properties of SWT promise superiority over the DWT based techniques [18,19]. In addition the diffusion part in shock filter reduces the sensitivity to noise and also enhance the visual quality of an image. We prefer shock filter over the bilateral filter [22] and the non-local means filter [23] due to its dual dominant operating modes for image enhancement as well as image denoising. The rest of the paper is organized as follows. A general description of shock filter is presented in Section 2. Our proposed algorithm is developed in Section 3. In Section 4 the experimental results are discussed. Finally, Section 5 concludes the paper.

2. Diffusion based shock filter

Shock filter has been proposed by Osher and Rudin [24] and it was proved successful in deblurring the images. The scope of shock filter is widely increased in several image enhancement and edge detection applications by regularizing the basic shock filter.

We first present a brief discussion of the shock filter. Consider an input image $X(m, n)$, with (m, n) as the spatial coordinates. Shock filter is viewed as tracing a time varying image $X(m, n, k)$, where k is the time index. These time variations can be captured from the first order time derivative

$$X_k = \frac{\partial X(m,n,k)}{\partial k}.$$

Here X_k is a function of m, n and k, which are dropped for notational simplicity. Given X_k we can evaluate $X(m, n, k)$ recursively starting with the input image as $X(m, n, k=0)$.

Mathematically, shock filter output X_k can be expressed as

$$X_k = -\operatorname{sign}(X_{\rho\rho})|\nabla X|, \tag{1}$$

where $X_{\rho\rho}$ is the second derivative of input image along ρ direction. ∇X is the corresponding gradient.

This filter is extremely sensitive to noise. The second derivative of the image is convolved with a Gaussian low pass filter to suppress the noise dominance.

$$X_k = -\operatorname{sign}(X_{\rho\rho} * G_\sigma)|\nabla X|. \tag{2}$$

The Gaussian filter with variance σ^2 is given by

$$G_\sigma(m,n) = \frac{1}{2\pi\sigma^2}\, e^{(-(m^2+n^2)}/2\sigma^2).$$

This simple convolution in Eq. (2) is inefficient to suppress the noise. The spikes due to noise still exist for moderate widths of Gaussian window. However, the convolution itself becomes costlier for larger widths. To overcome this, Gilboa et al. [25] optimized the filtered output by incorporating a weighted linear diffusion term to the original shock filter in Eq. (1)

$$X_k = -\operatorname{sign}(X_{\rho\rho})|\nabla X| + \lambda X_{\eta\eta}|\operatorname{sign}(X_\rho)|, \tag{3}$$

where $\lambda > 0$ is a diffusion weight and η is the direction perpendicular to ρ.

Filter presented in Eq. (3) gives improved results in image denoising and image enhancement as well. For small λ the diffusion term depletes to a great extent, resulting the process to be an edge preserving shock filter. However, for large λ image denoising plays a major role with a little edge enhancement. Thus the basic shock filter is slightly idle and not total-variation preserving (TVP). The diffusion term is multiplied by $|\operatorname{sign}(X_\rho)|$ to make it TVP. But this small manipulation will not give appreciable results. The magnitude of second derivative is employed to improve the robustness against noise. The basic shock filter in Eq. (3), is modified as:

$$X_k = -\frac{2}{\pi}\arctan(qX_{\rho\rho})|\nabla X| + \lambda X_{\eta\eta}, \tag{4}$$

where the parameter q facilitates control over the slope at zero crossings. The time index k can also be incorporated to reduce the errors in the evaluation of the gradients.

$$X_k = -\frac{2}{\pi}\arctan(qX_{\rho\rho}k)|\nabla X| + \lambda X_{\eta\eta}. \tag{5}$$

Furthermore, the above time dependent equation with zero potential and diffusion term can be generalized to complex values [26] as:

$$X_k = -\frac{2}{\pi}\arctan\left(q\,\mathrm{Im}\left(\frac{X}{\theta}\right)\right)|\nabla X| + \widetilde{\lambda}X_{\rho\rho} + \lambda X_{\eta\eta}, \tag{6}$$

where $\lambda \in \mathbb{R}$ and $\widetilde{\lambda} \in \mathbb{C}$ are the scalar diffusion weights and $\theta = arg(\widetilde{\lambda})$. We can also express the complex diffusion weight as $\widetilde{\lambda} = re^{j\theta}$.

The complex shock filter denoise the smoother parts of image as a result of diffusion equation and enhance edges of the image through the shock equation. This shock filter can be used to give improved results in image SR. For attempting the SR problem using subband decomposition, if we employ shock filter on the subbands, it ensures sharp edges as well as noise suppression.

3. Proposed SR reconstruction approach

We start with a discussion on a model for the LR image observation, followed by the proposed edge preserving SR algorithm. Later back projection algorithm developed from the LR image model, to improve the proposed SR technique is presented.

3.1. Modeling the LR image

The input LR image can be viewed as a blurred and downsampled observation of the HR image. The loss of high frequency details in an LR image is equivalent to blurring. From HR to LR sample density reduction implies downsampling. Bringing both operations together we can model the LR image observation Y as

$$Y = DBZ + V, \tag{7}$$

where Z is the original HR image, B is the blurring operator and D is the downsampling operator. Additive noise V takes care of modeling error and observation noise.

3.2. Specific SR reconstruction

To obtain the initial estimation of HR image Z_0, first the input LR image Y is enlarged by a factor s using bicubic interpolation. But simple interpolation techniques results in over smoothed images with inherent ringing and jagging artifacts. It will create major loss in the high frequency details. The final HR image has to be computed by further processing the initial estimate Z_0, by adding additional high frequency details. For this, subband decomposition is preferred [17]. Using the SWT the initial HR image Z_0 is decomposed into four different subbands, viz., approximation Z_A, horizontal Z_H,