Use of Bilateral Filters for Super-Resolution of Single Image with Iterations for Minimizing Reconstruction Errors
(バイラテラルフィルタを用いる再構成誤差最小化反復法による単一画像の超解像)

Chang Jian†, Kohei Inoue (member)†, and Kiichi Urahama†

Abstract The use of a nonlinear filter for super-resolution of a single image is presented. A nonlinear iterative solution procedure is derived from an approximate Newton method for minimizing errors in low-resolution images reconstructed from estimated high-resolution images. We use the bilateral filter for image smoothing in the reconstruction steps. Our method was shown to produce high-resolution images without halos around edges and with enhanced fine textures.

Key words: bilateral filter, super-resolution, halo, Newton method

1. Introduction

Algorithms for super-resolution of images are classified into interpolation techniques, reconstruction types and learning-based methods1) in which we consider the reconstruction type algorithms. While many techniques exploit redundancies in several images such as video data, we devote to single image super-resolution.

A popular super-resolution algorithm of reconstruction type called error back-projection method2) is a linear iterative solution method using a linear filter for smoothing images. We present, in this paper, a nonlinear iterative algorithm using the bilateral filter3) (BF in short) for minimizing reconstruction errors.

The BF was utilized for edge-detection in an edge-directed super-resolution method4). We use it, in this paper, directly for reconstructing images nonlinearly.

2. Reconstruction Error Minimization

Let the vector of pixel values in an input image be $d$. For color images, we decompose it into RGB components and we apply the following procedure to each component.

This $d$ is supposed to be shrunken after smoothing of an unknown large image $f$. We assume this smoothing is done with BF. Our goal is to reconstruct $f$ from $d$ of the input image.

We denote the operation of BF for $f$ by $F(f)$ of which pixel value $\hat{f}_{ij}$ is given by

$$\hat{f}_{ij} = \sum_{l=-p}^{p} \sum_{m=-p}^{p} w_{ijkl} f_{i+k,j+m} / \sum_{l=-p}^{p} \sum_{m=-p}^{p} w_{ijkl} \quad (1)$$

where $w_{ijkl} = e^{-α(δ^2 + m^2)} - \delta(f_{i,j} - f_{i,l,j+m})^2$.

We denote the process of shrinking the image $f$ be $D(f)$ which is, in this paper, down-sampling after bicubic interpolation of $f$. We denote the composite process $D(F(f))$ of shrinking after smoothing with BF for $f$ by $A(f)$. Then our purpose is to estimate $f$ from

$$d = A(f) \quad (2)$$

of which inverse $A^{-1}(d)$ is generally not unique.

As a candidate for $A^{-1}(d)$, we adopt $B(d) = S(U(d))$ where $U(d)$ is the bicubic interpolation after upsampling of $d$ and $S$ is the sharpening procedure:

$$S(f) = f + \delta(f - F(f)) \quad (3)$$

where $F$ is the same BF as used in the above-mentioned smoothing of images.

3. Iterative Solution Method

The error $d - A(f)$ in eq.(2) becomes $B(d - A(f))$ in reconstructed images. For computing $f$ minimizing this reconstruction error, we resort to an approximate Newton method for solving

$$B(d - A(f)) = 0 \quad (4)$$
The operator $S$ in $B(d) = S(U(d))$ is high-pass filter, hence $B(d - A(f)) = 0$ is attained only when $d - A(f) = 0$.

Generally, the Newton method for solving $s(x) = 0$ is to iteratively update $x$ by

$$x^{(\xi+1)} = x^{(\xi)} - t^{-1}(x^{(\xi)}) s(x^{(\xi)})$$  \hspace{1cm} (5)

which is derived from $s(x) \simeq s(x^{(\xi)}) + t(x^{(\xi)}) (x^{(\xi+1)} - x^{(\xi)}) = 0$ which is the Taylor series approximation of $s(x)$ around the $\xi$-th iterant $x^{(\xi)}$. In eq.(5), $t$ is the derivative of $S$, i.e. $t(x) = ds(x)/dx$.

For eq.(4), $t(x)$ is $dB(d - A(f))/df \cong -dB(A(f))/df$ because $B(d - A(f)) \cong B(d) - B(A(f))$ and $dB(d)/df = 0$. If $B(A(f))$ is identical mapping, i.e. $dB(A(f))/df = 1$, then $dB(d - A(f))/df \cong -1$ and the Newton iteration for eq.(4) is approximately reduced to

$$f^{(\xi+1)} = f^{(\xi)} + B(d - A(f^{(\xi)}))$$  \hspace{1cm} (6)

of which convergence is analyzed by the Taylor approximation: $B(d - A(f^{(\xi+1)})) \cong B(d - A(f^{(\xi)})) + dB(d - A(f^{(\xi)}))/df (f^{(\xi+1)} - f^{(\xi)}) = (I + dB(d - A(f))/df)B(d - A(f^{(\xi)}))$ from which we know the iteration converges if $\|I + dB(d - A(f))/df\| \leq 1$, i.e. $B(A(f))$ is close to identical mapping. We set the initial value as $f^{(0)} = U(d)$.

Eq.(6) is of form similar to the error back-projection iteration by Irani & Peleg\(^2\) which is an ad-hoc linear algorithm where $A(f)$ and $B(d)$ are linear operators. Contrastively eq.(6) is a nonlinear iterative algorithm derived from the Newton method. Our method includes a nonlinear filter BF and only local convergence is ensured in contrast to global convergence of the back-projection algorithm\(^3\).

Iteration of eq.(6) is expected to converge faster as closer $B(A(f))$ to the identity mapping. We therefore set its inverse mapping $A(B(d))$ as close as possible to the identity mapping by setting $\delta$ in eq.(3) to the value minimizing $\|d - A(B(d))\|$.  

4. Experiments

4.1 Setting of $\delta$

The value of $E(\delta) = ||d - A(B(d))||$ for the image in Fig.1(a) is plotted in Fig.1(b) where the horizontal axis is $\delta$. We set the parameters in BF as $\alpha = 0.01, \beta = 1$. The value of $E(\delta)$ takes its minimum at $\delta = 1$, hence we set $\delta = 1$ for this image.

![Fig. 1](image1.png)  
(a) input image

![Fig. 2](image2.png)  
(b) $E(\delta) = ||d - A(B(d))||$

![Fig. 3](image3.png)  
Output images.

![Fig. 4](image4.png)  
(a) bicubic

![Fig. 5](image5.png)  
(b) Dai et al.\(^5\)

Fig. 1 Variation in $E(\delta) = ||d - A(B(d))||$ with $\delta$.

Fig. 2 Variation in $\Delta^{(\xi)} = ||f^{(\xi+1)} - f^{(\xi)}||$.
4.4 Comparison with Other Methods

Fig.4(a) shows enlargement of Fig.1(a) with the bicubic interpolation. Edges and fine textures are blurred. Fig.4(b) is the result of the method by Dai et al. Fine textures are clearer in our method in Fig.3(a) than Fig.4(b).

Other example images are shown in Fig.5 and Fig.6. In these images, our method reconstructs fine textures, for instance, rugged sea bottom in Fig.5 and fine shades on the face. In these images, however, textures are over-enhanced in our results, hence the results of previous methods of moderately recovered textures seem adequate visually. This is due to our setting of parameters for these examples biased to strong enhancement of textures. Such comparison by our visual perception is subjective and only qualitative.

Their quantitative comparison is favored for verifying the performance of the methods. Unfortunately the original images are not given for all of the above examples. Therefore, we adopt popular images in Fig.7 as original ones and reduce their size for use as inputs to the methods. Fig.7(a) of size 512 × 512 is a standard example, while Fig.7(b) of size 512 × 512 contains high-frequency textures difficult for reconstruction, and Fig.7(c) of size 380×484 involves mainly low-frequencies easy to reconstruct. For these examples, we set parameters in our method to ordinary values for faithful reconstruction of images.

Among the previous methods, we implemented that of Kim & Kwon which is the only algorithm of which program is publicly opened. The PSNR and SSIM index between the original image and the reconstructed one is shown in tables from 1 to 3. The SSIM (Structural SIMilarity) index is a measure of image quality used sometimes for the evaluation of super-resolution methods. The SSIM takes the value between -1 and 1. The image quality is the best when SSIM=1. In these tables, “K.&K.” denotes the method by Kim and Kwon, “half” is the reconstruction from the size of half width and half height, and “quarter” is the recon-
struction from the size of quarter width and quarter height. While the PSNR and SSIM of K.&K. are often worse than the bicubic interpolation, the PSNR of the present method is highest. For the image "milkdrop", the SSIM of our method is lower than others, however it is still higher than 0.9. The PSNR and SSIM of several other methods are reported for the images lena and mandrill. Our method outperforms all of them. Thus our method works well for various images.

Reconstructed images are shown in Fig. 8 which are reconstructed from their quarter size. In these examples, the bicubic interpolation and the method by Kim and Kwon produce jaggy noises at pixels of very high frequency textures, e.g. the feather frill dangling from the hat in lena and fine face-hairs in mandrill. Our method can avoid such noisy artifacts.

5. Conclusion

A nonlinear iteration algorithm is derived from the Newton method for minimizing reconstruction errors in super-resolution of a single image. It is demonstrated with experiments that the present method can reconstruct images with lower errors than previous super-resolution methods. A subject of future research is improvement of anti-aliasing for reducing jaggy at strong edges.

[References]