J. Cent. South Univ. (2013) 20: 3180−3186
DOI: 10.1007/s11771-013-1842-y

Springer

Anisotropic fourth-order diffusion regularization for multiframe super-resolution reconstruction

HUANG Shu-ying(黄淑英)[1, 2], YANG Yong(杨勇)[3], WANG Guo-yu(王国宇)[1]

1. College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China;
2. School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330013, China;
3. School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013, China

Abstract: A novel regularization-based approach is presented for super-resolution reconstruction in order to achieve good tradeoff between noise removal and edge preservation. The method is developed by using L1 norm as data fidelity term and anisotropic fourth-order diffusion model as a regularization item to constrain the smoothness of the reconstructed images. To evaluate and prove the performance of the proposed method, series of experiments and comparisons with some existing methods including bi-cubic interpolation method and bilateral total variation method are carried out. Numerical results on synthetic data show that the PSNR improvement of the proposed method is approximately 1.0906 dB on average compared to bilateral total variation method, and the results on real videos indicate that the proposed algorithm is also effective in terms of removing visual artifacts and preserving edges in restored images.

Key words: super-resolution; anisotropic fourth-order diffusion; bilateral total variation; regularization

1 Introduction

Super-resolution (SR) image reconstruction refers to resolution enhancement of an image or a frame of a video sequence based on multiple LR observed frames. In many practical applications, such as remote sensing [1], video surveillance [2], medical diagnostics [3], the demand for high-resolution images is gradually increasing. However, limited by physical constraints, the quality of image resolution can not meet the increasing demands. Therefore, it is essential to find an effective way to expand the resolution of low-resolution (LR) images.

To obtain a high-resolution image from a sequence of low-resolution images, researchers have explored many approaches to produce HR images in the last three decades. These methods are roughly categorized into two classes, namely, frequency domain methods and spatial domain methods. In the frequency domain, TSAI and HUANG [4] first demonstrated the reconstruction of improved resolution images by estimating the relative shifts between observations. Based on this algorithm, a series of improved SR reconstruction algorithms have been proposed. However, although frequency domain methods are computationally attractive, they have some limitations for they are not able to incorporate any spatial domain prior knowledge in their formulation [5]. Therefore, many spatial domain methods [6−9] have been successfully developed to overcome the aforementioned drawbacks. Among all spatial domain approaches, the regularization-based methods are effective to solve the multi-frame SR reconstruction problem. Because SR reconstruction is an ill-posed inverse problem, their common point is to obtain a stable solution by integrating a priori knowledge into the process of SR reconstruction. In the past decades, the partial differential equations (PDEs) show a great potential to achieve a good trade-off between noise removal and edge preservation. The total variation (TV) regularization methods and its variants have also become popular in image restoration [10−12] because of its ability to preserve edge information. However, the TV regularization model as the second order PDE also has some shortcomings. One of them is that the TV model favors a piecewise constant solution, which causes the

Foundation item: Projects(60963012, 61262034) supported by the National Natural Science Foundation of China; Project(211087) supported by the Key Project of Ministry of Education of China; Projects(2010GZS0052, 20114BAB211020) supported by the Natural Science Foundation of Jiangxi Province, China
Received date: 2012−06−20; **Accepted date:** 2012−09−20
Corresponding author: YANG Yong, Professor; Tel: +86−13803524350; E-mail: greatyangy@126.com

staircase effect resulting in the flat regions of the image. In order to overcome such drawbacks, fourth-order PDEs have been suggested for image restoration [13]. Although fourth-order diffusion methods reduce blocky effects on the restored image, they are more severely prone to the over-smoothness due to higher strength of smoothness of the filters compared to the second order filters. Another recently introduced anisotropic fourth-order diffusion (AFOD) filter by HAJIABOLI [14] can obtain a good edge preservation capability compared to the other diffusion filters by choosing suitable diffusivity functions to unevenly control the strength of the diffusion on the directions of the level set and gradient. Inspired by this, we propose a novel regularization-based SR reconstruction algorithm, which uses AFOD as a regularization item to constrain the smoothness of the reconstructed images and use L1 norm as data fidelity term. The motivation of our method is to achieve good tradeoff between noise removal and edge preservation during the procedure of reconstruction. In order to test its practical potential, we have applied our method to synthetic images and real video images and also compared our method with some related works.

2 Mathematical background

2.1 Observation model

In SR image reconstruction, it is necessary to set up a suitable imaging degradation model, which is to relate the desired reference high-resolution image with all the observed low-resolution images. Suppose there are p LR images Y_1, Y_2, ⋯, Y_p from the HR image u, where the size of the k-th LR image is $M\times M$, and the size of the HR image is $rM\times rM$, where the parameter r is the down-sampling factor. The relationship between an ideal HR image and an LR image can be described as [15]

$$\boldsymbol{Y}_k = \boldsymbol{D}_k\boldsymbol{B}_k\boldsymbol{E}_k u + \boldsymbol{n}_k \quad \text{for } 1\le k\le p \tag{1}$$

where $\boldsymbol{E}_k$ is the motion matrix for modeling the motion degradation, $\boldsymbol{B}_k$ is the blurring matrix for representing the point spread function (PSF) of the camera sensor, $\boldsymbol{D}_k$ is a down-sampling matrix for presenting the loss of resolution in the observed images, $\boldsymbol{n}_k$ is the normally distributed additive noise.

By stacking the vector equations from all the frames, we obtain a matrix vector formula:

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix} = \begin{bmatrix} D_1B_1E_1 \\ \vdots \\ D_kB_kE_k \end{bmatrix} u + \begin{bmatrix} n_1 \\ \vdots \\ n_k \end{bmatrix} \tag{2}$$

which can be also represented by

$$\boldsymbol{Y}=\boldsymbol{A}U+\boldsymbol{N} \tag{3}$$

2.2 Fourth-order PDE

The nonlinear diffusion filters are evolutionary processes. Many researchers have applied the ideas of PERONA and MALIK [16] to higher order equations in an effort to reduce staircasing effects. The second order PDE is given as

$$\frac{\partial u}{\partial t} = \nabla\left(g\left(\left\|\nabla u\right\|\right)\nabla u\right) \tag{4}$$

where u is the image intensity function, $g(\cdot)$ is a nonlinear diffusivity function which would inhibit diffusion across edges, t is the evolution time, and $\|\cdot\|$ is used for mathematical notation of Euclidean norm. One of the diffusivity functions suggested by PERONA and MALIK is:

$$g\left(\left\|\nabla u\right\|\right) = \frac{k^2}{k^2+\left\|\nabla u\right\|^2} \quad \text{for a constant } k>0. \tag{5}$$

In order to resolve formation of staircase artifacts on the ramp edges that are widely seen in images processed by the second order PDE model, a fourth order PDE for noise removal was proposed by YOU and KAVEH [13]. The energy function is defined as

$$E(u) = \iint_{\Omega} f\left(\left|\nabla u^2\right|\right)\mathrm{d}x\mathrm{d}y \tag{6}$$

where Ω is the image support, and $f(\cdot)\ge 0$ is an increasing function associated with the diffusion coefficient as

$$f'(s) = sg(s) \tag{7}$$

The corresponding Euler equation of Eq. (6) can be solved through the following gradient descent procedure:

$$\frac{\partial u}{\partial t} = -\nabla^2\left[f'\left(\left|\nabla^2 u\right|\right)\frac{\nabla^2 u}{\left|\nabla^2 u\right|}\right] = -\nabla^2\left[g\left(\left|\nabla^2 u\right|\right)\nabla u^2\right] \tag{8}$$

2.3 Anisotropic diffusion model

The extent of the diffusion of the isotropic filters is controlled by the diffusivity function regardless of the orientation of the edges. The anisotropic diffusion filters with good edge preservation capability control the diffusion rate with respect to the direction of the local features. The PDE of this filter is given by

$$\frac{\partial u}{\partial t} = -g_1(g_2 u_{\eta\eta} + g_3 u_{\xi\xi}) \tag{9}$$

where g_1, g_2 and g_3 are different diffusivity functions and $u_{\eta\eta}$ and $u_{\xi\xi}$ are the second-order directional derivatives in the directions of η and ξ. Let η denote the gradient direction, and ξ denote the direction perpendicular to the gradient, also known as the direction of level set. The definitions are given by

$$\eta=\frac{[u_x,u_y]}{\sqrt{u_x^2+u_y^2}} \text{ and } \xi=\frac{[-u_y,u_x]}{\sqrt{u_x^2+u_y^2}} \tag{10}$$

The second-order directional derivatives are defined as

$$u_{\eta\eta}=\frac{u_{xx}u_x^2+2u_xu_yu_{xy}+u_{yy}u_y^2}{u_x^2+u_y^2} \tag{11}$$

$$u_{\xi\xi}=\frac{u_{xx}u_x^2-2u_xu_yu_{xy}+u_{yy}u_y^2}{u_x^2+u_y^2} \tag{12}$$

Based on the model, HAJIABOLI [14] proposed a fourth-order nonlinear diffusion filter for image denoising. The filter can better preserve edge by unevenly controlling the diffusion on the directions of level set and gradient. The definition of the filer is

$$\frac{\partial u}{\partial t}=-\nabla^2\left(g_1\left(g_2u_{\eta\eta}+g_3u_{\xi\xi}\right)\right) \tag{13}$$

In the model, g_1 controls total amount of diffusion, g_2 and g_3 control the uneven diffusion in the direction of η and ξ. The most commonly used diffusivity function in fourth-order diffusion filters is the one in Eq. (5) as $g(s)$, where s is the modulus of the derivative of the image. A suitable choice for these diffusivity functions in terms of overall computational cost of the filter is to select as $g_1(s)=g_2(s)=g(\|\nabla u\|)$ and $g_3(s)=1$. Therefore, the fourth-order nonlinear diffusion filter can be written as

$$\frac{\partial u}{\partial t}=-\nabla^2\left(g(\|\nabla u\|)^2u_{\eta\eta}+g(\|\nabla u\|)u_{\xi\xi}\right) \tag{14}$$

It has been proved in the literature that the filter compared to the second order filter of PERONA and MALIK has the ability of preserving the ramp edges. In the next section, the fourth-order nonlinear diffusion filter will be applied as the regularization term to suppressing the noise in the super-resolution reconstruction.

3 Proposed AFOD based super-resolution algorithm

Based on the observation model Eq. (1), the SR image can be estimated by minimizing a least-square cost function, such as

$$\hat{u}=\underset{u}{\text{Arg min}}\left[\sum_{k=1}^{p}\left\|\boldsymbol{D}_k\boldsymbol{B}_k\boldsymbol{E}_ku-\boldsymbol{Y}_k\right\|^2\right] \tag{15}$$

Due to the ill-posed nature of SR reconstruction, the solution of Eq. (15) is generally not unique. In order to obtain more desirable SR results, the ill-posed problem should be stabilized to become well-posed. A regularized approach using the image prior information of the desired HR image can compensate for the missing measurement information, and make the inverse problem well-posed. The corresponding HR estimation can be calculated by the following cost function minimization:

$$\hat{u}=\arg\min_u\left[\sum_{k=1}^{p}\left\|D_kB_kE_ku-Y_k\right\|^2+\lambda\gamma(u)\right] \tag{16}$$

where the first term is the data fidelity term, which measures the reconstruction error to ensure that pixels in the reconstructed HR image are close to real values; the second term as the regularization term is to control the smoothness of the reconstructed HR image and λ is the regularization parameter which provides a tradeoff between the data fidelity term and the regularization term.

In this work, to better remove visual artifacts and preserve image edges information, we employ the anisotropic diffusion fourth-order PDE as the regularization term for the SR image reconstruction. Therefore, by combining Eq. (14) and the Euler-Lagrange equation of Eq. (16), the AFOD based super-resolution reconstruction can be yielded as

$$\nabla E(u)=\sum_{k=1}^{N}\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{D}_k^{\mathrm{T}}\text{sign}(\boldsymbol{D}_k\boldsymbol{B}_k\boldsymbol{E}_ku-\boldsymbol{Y}_k)+\lambda\left[-\nabla^2\left(g(\|\nabla u\|)^2u_{\eta\eta}+g(\|\nabla u\|)u_{\xi\xi}\right)\right]=0 \tag{17}$$

We use steepest descent to find the solution to this minimization problem.

$$u^{n+1}=u^n-\beta\nabla E(u)=u^n-\beta\{\sum_{k=1}^{N}\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{B}_k^{\mathrm{T}}\boldsymbol{D}_k^{\mathrm{T}}\text{sign}\cdot(\boldsymbol{D}_k\boldsymbol{B}_k\boldsymbol{E}_ku^n-\boldsymbol{Y}_k)+\lambda[\nabla^2(g(\|\nabla u^n\|)^2u_{\eta\eta}^n+g(\|\nabla u^n\|)u_{\xi\xi}^n)]\} \tag{18}$$

where n is the iteration number, and β is a scalar defining the step size in the direction of the gradient. The matrices $\boldsymbol{D}_k$, $\boldsymbol{B}_k$ and $\boldsymbol{E}_k$ model the image formation process, and their implementation is simply the subsampling, blurring and image warping, respectively. $\boldsymbol{D}_k^{\mathrm{T}}$ is implemented by upsampling the image without interpolation, i.e., zero padding. $\boldsymbol{B}_k^{\mathrm{T}}$ is implemented by convolution with the flipped kernel, i.e. if H is the imaging blur kernel, then the flipped kernel is the matrix $\boldsymbol{K}$ such that $\forall i,j,K(i,j)=H(-i,-j)$. If $\boldsymbol{E}_k^{\mathrm{T}}$ is implemented by backward warping, then E_k^{T} should be the forward warping of the inverse motion.

4 Experimental results

Three reconstructed results of two sets of simulated data derived from a single image and one set of real video image sequences are presented. The performance of the proposed algorithm is compared to those of bi-cubic interpolation (BCI) and the bilateral total

variation (BTV) SR method. To better evaluate the performance of the algorithm, we use the following peak signal-to-noise ratio (PSNR) defined as a quantitative measure in the synthetic images experiments.

$$f_{\mathrm{PSNR}} = 10\lg\left(\frac{255^2 \times rM \times rM}{\|\hat{u}-u\|^2}\right) \tag{19}$$

where $rM \times rM$ is the total number of pixels in the HR image, and $\hat{u}$ and u represent the reconstructed HR image and original image, respectively.

4.1 Simulated data

Two simulated data sets were used to test the effectivity and robustness of the proposed method: the "Butterfly" image of size 366×243 pixels and the "Cameraman" image of size 256×256 pixels, respectively, as shown in Fig. 1(a) and Fig. 2(a). For the first simulated experiment, the "Butterfly" image was shifted in the vertical direction by 0, 1 or 2 pixels and in the horizontal direction by 0, 1, or 2 pixels to produce eight LR images. Then, each shifted image was convolved with a PSF kernel of 3×3 window size with a standard deviation σ=0.5. The resulting image was down-sampled by a factor of 3 in each direction. Lastly, zero-mean Gaussian-noise with 0.005 variance was added to the LR images. Figure 1(b) shows one of the observed low-resolution degraded images. The reconstruction results of the "Butterfly" image are shown in Figs. 1(c)−(e), respectively.

For the second simulated experiment, the "Cameraman" image was shifted in the vertical direction by 0, 1 or 2 pixels and in the horizontal direction by 0, 1, or 2 pixels to produce eight LR images. Then, each shifted image was convolved with a PSF kernel of 3×3 window size with a standard deviation σ=1. The resulting image was down-sampled by a factor of 2 in each direction. Lastly, zero-mean Gaussian-noise with 0.01 variance was added to the LR images. Figure 2(b) shows one of the observed low-resolution degraded images. The reconstruction results of the "Cameraman" image are shown in Figs. 2(c)−(e), respectively.

Besides the above experiments, to better verify the robustness of the proposed method, we also tested it on other popular widely used test images, i.e., "Lena", "Barbara", and "Boat" images. The total quantitative comparisons of the reconstruction results using PSNR are listed Table 1.

From the comparison of the results, we can find that

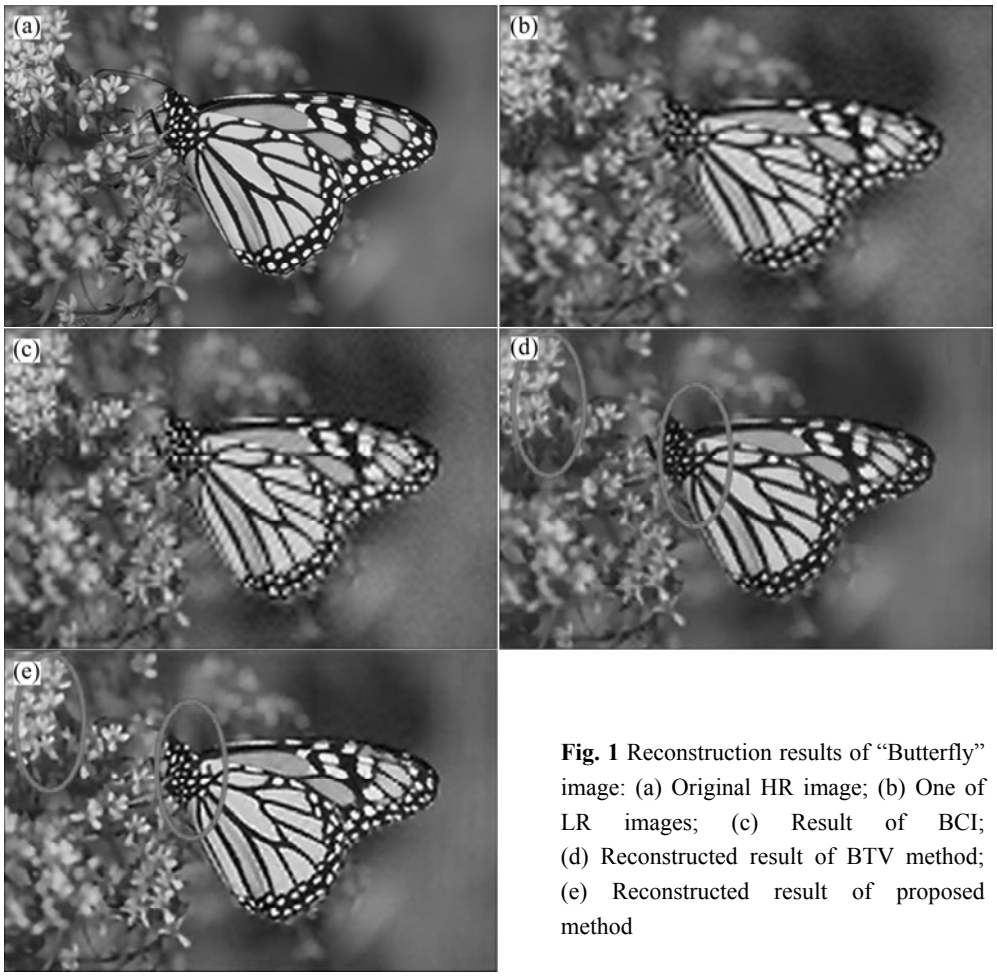

Fig. 1 Reconstruction results of "Butterfly" image: (a) Original HR image; (b) One of LR images; (c) Result of BCI; (d) Reconstructed result of BTV method; (e) Reconstructed result of proposed method

Fig. 2 Reconstruction results of "Cameraman" image: (a) Original HR image; (b) One of LR images; (c) Result of BCI; (d) Reconstructed result of BTV method; (e) Reconstructed result of proposed method

Table 1 Quantitative evaluation results of PSNR values for five test images

Sequence	BCI	BTV	Proposed method
Butterfly	22.204 5	23.323 7	25.197 6
Cameraman	24.277 2	24.560 9	25.724 8
Lena	26.293 3	28.301 4	28.611 0
Barbara	23.424 4	23.450 8	23.631 9
Boat	25.953 3	26.515 6	27.349 7

the result of BCI is the poorest in the three methods, which confirms the promise of the SR technology. It also can be seen that although the AFOD reconstruction method and BTV method provide better results than BCI method, the AFOD reconstruction method presents more details than the BTV method, for examples, in Fig.1, the flower of the image and the head of the Butterfly as shown in the marked ellipse areas of Fig. 1(d) and Fig. 1(e), respectively, and in Fig.2, the edges of the man

and the door of the house as shown in the marked circle areas of Fig. 2(d) and Fig. 2(e), respectively. By comparison, we also can easily find that the results of the proposed method are close to the original HR image and our method provides the highest PSNR value, which indicates that the proposed method can provide good representation of the original HR image, both in terms of gray values and structure.

4.2 Real video data

To further illustrate the performance of the proposed method, we tested it on one real video data set. The real data "disk" video sequence was obtained from the Multidimensional Signal Processing(MDSP) Research Group of UCSC [17], consisting of 20 frames with the size of 49×57 pixels. The registration approach presented in Ref. [18] was used as the motion estimation method. Here, we increase the resolution by a factor of 4 to LR images. The camera blur kernel is a 7×7 Gaussian kernel with a standard deviation 2.0.

Figure 3(a) shows one of the LR frames of the video, the BCI and BTV results are shown in Figs. 3(b) and (c), respectively, and Fig. 3(d) shows the result of the proposed method. From Figs.3 (a)−(d), we can see that the resolution has certainly increased after using the SR technique compared with the results of the BCI interpolated results. Of the two SR methods, the proposed AFOD method provides a clearer result than the BTV. It should be noted that since the real data has no ground truth, the above PSNR metric can not be used in this case. Therefore, the reconstructed results can only be assessed by subjective evaluation as in most literature.

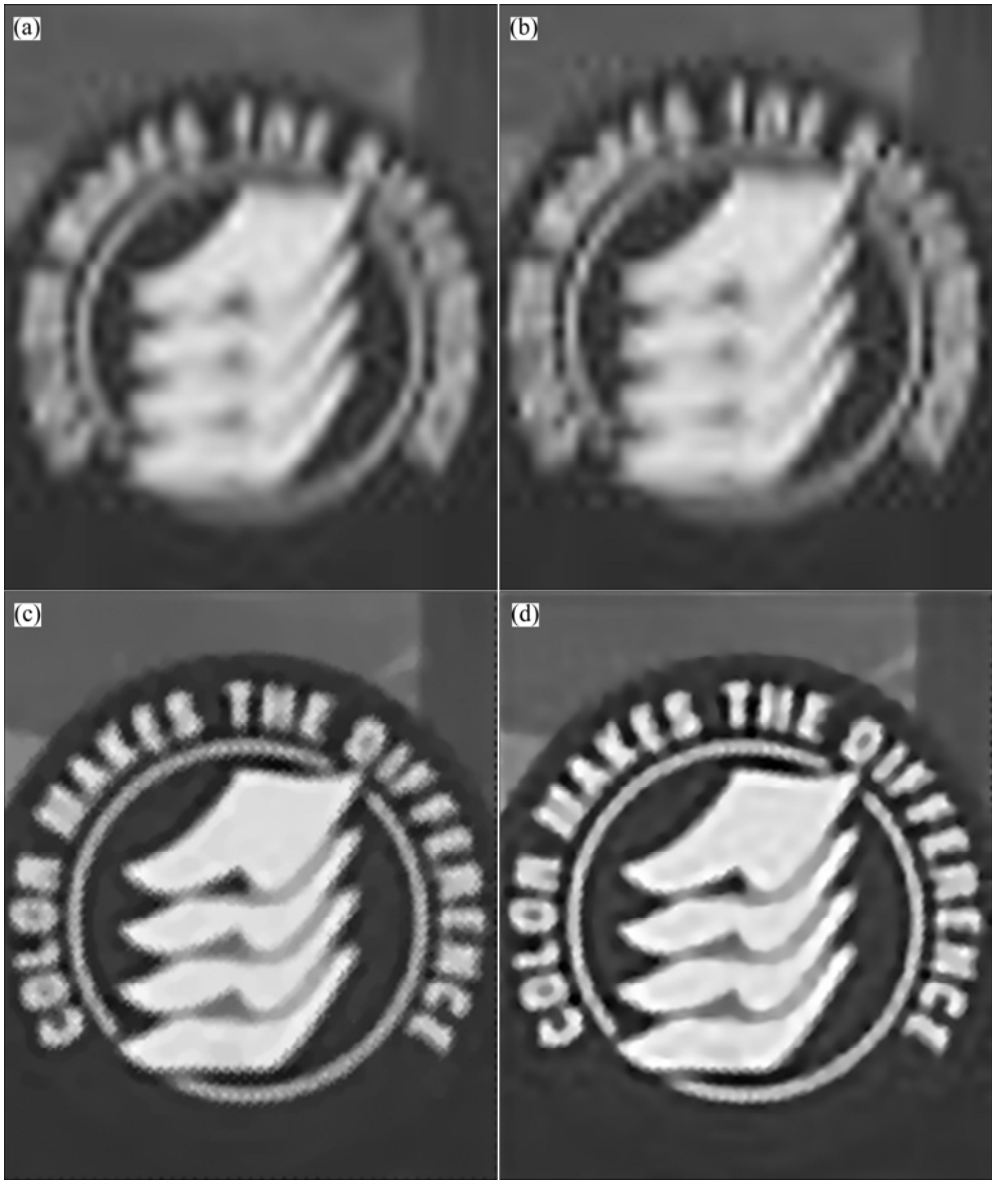

Fig. 3 Reconstruction results of "disk" image: (a) One of LR images; (b) Result of BCI; (c) Reconstructed result of BTV method; (d) Reconstructed result of proposed method

5 Conclusions

1) In order to overcome the staircase effect and the loss of fine details of the second order PDE based image reconstruction algorithm, a novel super-resolution method is presented by constructing a new energy function, which is composed of L_1 norm minimization and robust regularization based on anisotropic diffusion fourth-order PDE.

2) The proposed algorithm is tested on different series of simulated images and real video sequences and also compared with other three existing methods. The proposed approach is able to enhance the quality of a set of noisy blurred images and efficiently remove outliers, which can achieve resultant images with sharp edges and minimal artifacts. The experimental results show that on synthetic data, the algorithm outperforms BCI interpolation by 2.090 5 dB and BTV algorithm by 1.090 6 dB in PSNR values, and in all cases, the proposed algorithm can achieve better visual results in terms of noise suppression and detail preservation.

References

[1] VEGA M, MATEOS J, MOLINA R, KATSAGGELOS A.K. Super-resolution of multispectral images [J]. The Computer Journal, 2009, 52(1): 153−167.

[2] ZHANG L P, ZHANG H Y, SHEN H F, LI P X. A super-resolution reconstruction algorithm for surveillance images [J]. Signal Processing, 2010, 90(3): 848−859.

[3] CARMI E, LIU S, ALON N, FIAT A, FIAT D. Resolution enhancement in MR [J]. Magnetic Resonance Imaging, 2006, 24(2): 133−154.

[4] TSAI R Y, HUANG T S. Multiple frame image restoration and registration [J]. Advances in Computer Vision and Image Processing, 1984, 1(2): 317−339.

[5] PARK S C, PARK M K, KANG M G. Super-resolution image reconstruction: A technical overview [J]. IEEE Signal Processing Magazine, 2003, 20(3): 21−36.

[6] PROTTER M, ELAD M, TAKEDA H, MILANFAR P. Generalizing the nonlocal-means to super-resolution reconstruction [J]. IEEE Transactions on Image Processing, 2009, 18(1): 36−51.

[7] XIONG Z W, SUN X Y, WU F. Robust web image/video super-resolution [J]. IEEE Transactions on Image Processing, 2010, 19(8): 2017−2028.

[8] YUAN Q, ZHANG L, SHEN H, LI P. Adaptive multiple-frame image super-resolution based on U-curve [J]. IEEE Trans Image Process, 2010, 19(12): 3157−3170.

[9] DONG W, ZHANG L, SHI G, WU X. Image deblurring and supper-resolution by adaptive sparse domain selection and adaptive regularization [J]. IEEE Transactions on Image Processing, 2011, 20(7): 1838−1857.

[10] NG M K, SHEN H, LAM E Y, ZHANG L. A total variation regularization based super-resolution reconstruction algorithm for digital video [J]. EURASIP Journal on Advances in Signal Processing, 2007, 74585.

[11] CHEN Q, MONTESINOS P, SUN Q, HENG P, XIA D. Adaptive total variation denoising based on difference curvature [J]. Image Vis Comput, 2010, 28(3): 298−306.

[12] MARQUINA A, OSHER S. Image super-resolution by TV-regularization and bregman iteration [J]. Journal of Scientific Computing, 2008, 37(3): 367−382.

[13] YOU Y L, KAVEH M. Fourth-order partial differential equations for noise removal [J]. IEEE Transactions on Image Processing, 2000, 9(10): 1723−1730.

[14] HAJIABOLI M R. An anisotropic fourth-order diffusion filter for image noise removal [J]. International Journal of Computer Vision, 2011, 92(2): 177−191.

[15] FARSIU S, ROBINSON M D, ELAD M, MILANFAR P. Fast and robust multi-frame super-resolution [J]. IEEE Transactions on Image Processing, 2004, 13(10): 1327−1344.

[16] PERONA P, MALIK J. Scale-space and edge detection using anisotropic diffusion [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(7): 629−639.

[17] MDSP. Super-resolution and Demosaicing Datasets [EB/OL]. [2007−01−01]. http://users.soe.ucsc.edu/~milanfar/software/sr-tasets.html

[18] SCHULTZ R R, MENG L, STEVENSON R L. Subpixel motion estimation for super-resolution image sequence enhancement [J]. Journal of Visual Communication and Image Representation, 1998, 9(1): 38−50.

(Edited by DENG Lü-xiang)