A robust multiframe super-resolution algorithm based on half-quadratic estimation with modified BTV regularization

Xueying Zeng a, Lihua Yang b,∗

a School of Mathematical Sciences, Ocean University of China, Qingdao, 266100, PR China
b Guangdong Province Key Laboratory of Computational Science, School of Mathematics and Computational Science, Sun Yat-sen University, Guangzhou, 510275, PR China

1. Introduction

Multiframe super-resolution (SR) reconstruction aims to combine a set of low-resolution (LR) degraded images portraying slightly different views of the same scene to reconstruct a high-resolution (HR) image of that scene. The idea is to improve the details of the reconstructed HR image by exploiting the additional information available due to the subpixel motion among the captured LR images. Problems motivating SR arise in a number of image application fields, such as remote sensing, medical diagnosis and military information acquisition.

The multiframe image super-resolution problem was first addressed in [1]. Since then, numerous approaches have been proposed and studied for solving SR problem. Regularization method is a widely used approach to solve SR problem due to its ill-posed nature. Recent efforts based on the regularization framework for the SR problem are the works in [2–17]. Regularized SR reconstruction obtains an HR image by minimizing an objective function consisting of a fidelity term and a regularization term. The fidelity term measures the closeness between the estimated HR image and the captured LR images. The regularization term is utilized to regularize the problem and achieve the stable solution. Regularization SR reconstruction can be also viewed as a maximum a posteriori approach, because the fidelity term can be matched to those of

© 2012 Elsevier Inc. All rights reserved.

http://dx.doi.org/10.1016/j.dsp.2012.06.013
the Gaussian error norm [20]. El-Yamany et al. developed an adaptive robust M-estimation scheme without regularization, where the Lorentzian error norm is used [21,19]. Other robust M-estimators include Hampel, Andrew’s Sine, Leclerc and Tukey’s Biweight estimators, Patanavijit et al. presented a series of regularized SR algorithms by using these robust estimators and the Lorentzian error norm in both the fidelity and regularization terms [11,22–25]. Although these robust estimators increase the robustness against outliers, they cause the minimizing functional to be unconvex. Therefore, even the simplest steepest descent method used in these papers cannot guarantee its convergence to the solution of the minimization problem. Moreover, due to the nonconvexity of these robust estimators, some sophisticated optimization methods such as conjugate gradient algorithm are not easy to accurately used for solving their resulting SR problems. This motivates us to find an estimator combining both advantages of the $L_1$ and $L_2$ norms in the class of convex functions.

In this paper, we propose a new robust norm for the regularized SR problem both in the fidelity and regularization terms. This norm can adaptively mimic $L_1$ and $L_2$ norms. Using it in fidelity term, an adaptive strategy was developed to effectively deal with different accuracy levels of the estimated model parameters among LR frames. This strategy serves to adaptively weight LR frames according to their reliability, and hence can automatically suppress the contribution of those frames that suffer from inaccuracies in their estimated parameters to the final HR image. The proposed regularization term, called bilateral edge-preserving regularizer, captures the correlation between two pixels by both their spatial distance and photometric distance. Using the proposed norm to penalize the photometric distance (i.e. gradient magnitudes) can preserve large gradients corresponding to edges, while smooth small gradients usually are effects of noise. The formed objective function is twice continuously differentiable and strictly convex, and hence the gradient-based optimization technique can find the unique optimal super-resolution image.

This paper is organized as follows. We propose the imaging model of SR in the next section. Section 3 proposes the new robust norm and develops the strategy to adaptively weight LR frames. In Section 4, we describe the proposed bilateral edge-preserving regularizer and present our regularized SR algorithm. In Section 5, we present experiments with synthetic and real data to justify our algorithm. Finally, we conclude this paper.

2. Image observation model

The first step to analyze the SR image reconstruction problem is to formulate an observation model that relates the original HR image to the observed LR images. Typically, the imaging process involves warping followed by blurring and downsampling to generate LR images from an original HR image [18]. Let the underlying HR image be written in lexicographical notation as the vector $X = [x_1, x_2, \ldots, x_N]^T$, where $N = L_1N_1 \times L_2N_2$ is the HR image size. Letting $L_1$ and $L_2$ denote the downsampling factors in the horizontal and vertical directions, respectively, each observed LR image is of size $N_1 \times N_2$. Thus, the LR image can be denoted as $Y_k = [y_{k1,1}, y_{k2,2}, \ldots, y_{kM,M}]^T$, where $k = 1, 2, \ldots, p$, and $M = N_1 \times N_2$, with $p$ being the number of the LR images. Assuming that each LR image is corrupted by additive noise, the observation model can be represented as [2,7,18]

$$Y_k = DB_{k}M_kX + n_k \quad (1)$$

where $M_k$ is a warp matrix of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, $B_k$ represents an $L_1N_1L_2N_2 \times L_1N_1L_2N_2$ blur matrix, $D$ is an $N_1N_2 \times L_1N_1L_2N_2$ downsampling matrix, and $n_k$ represents the $N_1N_2 \times 1$ noise vector, and usually is assumed to be zero mean Gaussian noise.

The relationship between an HR image and LR observations can be also rewritten as

$$y_{k,m} = \sum_{r=1}^{N} c_{r,m}^{k}x_{r} + n_{k,m} \quad (2)$$

where $k = 1, 2, \ldots, p$ and $m = 1, 2, \ldots, M$, $c_{r,m}^{k}$ represents the “contribution” (including warping, blurring, and downsampling) of the $r$-th HR pixel $x_r$ to the $m$-th LR observed pixel of the $k$-th frame.

3. The proposed adaptive fidelity term

3.1. The $M$-estimation framework

$M$-estimation [26,27], in the SR context, is based on the minimization of a cost function which measures the residual between the captured LR images and the estimated HR image. We write the definition of $M$-estimation as the following minimization problem

$$\hat{X} = \arg\min_{X} \sum_{k=1}^{p} \rho(DB_{k}M_kX - Y_k)$$

$$= \arg\min_{X} \sum_{k=1}^{p} \sum_{m=1}^{M} \rho(e_{k,m}) \quad (3)$$

or by an implicit equation

$$\sum_{k=1}^{p} \sum_{m=1}^{M} \psi(e_{k,m}) = 0$$

where $e_{k,m} = \sum_{r=1}^{N} c_{r,m}^{k}x_{r} - y_{k,m}$, with $c_{r,m}^{k}$, $x_{r}$, $y_{k,m}$ the same as in Eq. (2). $\rho(x)$ is an even symmetric positive function that has a unique minimizer at $x = 0$. $\psi(e_{k,m}) = (\frac{1}{\rho(x)})\rho'(e_{k,m})$ is the first derivative of $\rho(e_{k,m})$ with respect to $x$.

The function $\rho$ in Eq. (3) is called an M-estimator because it corresponds to the ML type estimation [28]. Assuming the zero mean Gaussian noise model in Eq. (1), the ML estimation of HR image can be achieved when $\rho(x) = \frac{x^2}{2}$

$$\hat{X} = \arg\min_{X} \sum_{k=1}^{p} \frac{1}{2} ||DB_{k}M_kX - Y_k||_2^2$$

$$= \arg\min_{X} \sum_{k=1}^{p} \sum_{m=1}^{M} \frac{1}{2} e_{k,m}^2 \quad (4)$$

However, because of the $L_2$ error norm in the cost function in Eq. (4), the solution exhibits a poor performance in the presence of intensity outliers or large registration errors [7]. The reason for the nonrobustness of the $L_2$ error norm lies in its linear influence function. The influence function $\psi(x)$ of an M-estimator $\rho(x)$ is defined by $\psi(x) = x\rho'(x)$. This function characterizes the bias that a particular measurement has on the solution [26]. As shown in Fig. 1(b), $\psi_2(x) = x$ is linearly proportional to the error, assigning larger weights to larger errors, and hence outliers can therefore bias the solution to an erroneous value.

Farsiu et al. made a progress by proposing the use of the $L_1$ error norm as a robust alternative to the $L_2$ error norm [7], and the reconstructed HR image is solved by

$$\hat{X} = \arg\min_{X} \sum_{k=1}^{p} ||DB_{k}M_kX - Y_k||_1$$

$$= \arg\min_{X} \sum_{k=1}^{p} \sum_{m=1}^{M} |e_{k,m}| \quad (5)$$
The $L_1$ error norm, where $\psi_1(x) = \text{sign}(x)$, is not sensitive to outliers by assigning the same weights (+1 or −1) to all errors (small and large). However, the constant-valued influence does not distinguish between large errors often caused by outliers or registration errors and small errors corresponding to observation noise. The $L_1$ norm therefore produces an estimator with higher variance than $L_2$ norm. As a result, when the residual errors are still approximately Gaussian [3], $L_1$ norm does not perform as well as $L_2$ norm. Moreover, $\rho(x) = |x|$ is not differentiable at zero, the gradient descent algorithm used in [7] will introduce numerical instability in the iteration process [29].

To increase the robustness to large errors while still being effective to small approximately Gaussian errors, Patanavijit et al. used Huber function in fidelity term for measuring the difference between the estimated HR image and the captured LR images [8]. As shown in Fig. 1, Huber estimator can combine the behavior of $L_2$ norm when the errors are small while maintaining the $L_1$ norm’s reduced sensitivity to larger errors. To increase robustness further, some redescending M-estimators have been explored recently in multiframe SR reconstruction. Pham et al. used the robust Gaussian error norm in the cost function in Eq. (3) without regularization [20]. El-Yamany et al. developed an adaptive M-estimation scheme without regularization, where the robust Lorentzian error norm is used in Eq. (3) [21,19]. Patanavijit et al. used the Lorentzian, Hample, Andrew’s Sine, Leclerc and Tukey’s Biweight error norms in both the fidelity and regularization terms to present a series of regularization SR algorithms [11,22–25]. Such redescending M-estimators (Gaussian, Lorentzian, Hample, Andrew’s Sine, Leclerc and Tukey’s Biweight) all have two important properties:

(i) For values smaller than a threshold, their influence functions are approximately linear and they perform like $L_2$ norm.

(ii) For values larger than the threshold, their influence functions tend to zero and therefore can effectively suppress outliers.

In this sense, these robust M-estimators are expected to have similar performance in SR problems. The thing is to select a suitable threshold for each estimator. However, they cause the cost function in Eq. (3) to be nonconvex, and hence the used steepest descent optimization technique in these papers can only find a local minimizer. Which local minimizer is obtained sensitively depends on the choice of the iterative initial [28,30]. Moreover, due to their nonconvexities, many sophisticated convex optimization algorithms cannot be used to solve the resulting minimization problems. As two examples of robust estimators, we list the definitions of the Leclerc and Lorentzian error norms below

$$\rho_{\text{Lec}}(x, T) = 1 - e^{-\frac{x^2}{T^2}}, \quad \rho_{\text{Lor}}(x, T) = \log \left( 1 + \frac{x^2}{2T^2} \right)$$

where parameter $T$ controls the outlier threshold. Their influence functions are defined as

$$\psi_{\text{Lec}}(x, T) = \frac{2x}{T^2} e^{-\frac{x^2}{T^2}}, \quad \psi_{\text{Lor}}(x, T) = \frac{2x}{2T^2 + x^2}$$

The Leclerc error norm with $T = 50$, Lorentzian error norm with $T = 50$ and their corresponding influence functions are shown in Figs. 1 (a) and (b), respectively, we can clearly see their above two properties.

3.2. Adaptive fidelity term based on a new M-estimator

To automatically reduce the contribution of large errors and average the small approximately Gaussian errors, we hope to choose an estimator, whose influence function behaves like linearly proportional to small errors and then gradually becomes a constant to be robust against large errors. This estimator should be also strictly convex and twice continuously differentiable to make the corresponding minimization problem well posed. The following used error norm satisfies these requirements

$$\rho(x, a) = a \sqrt{a^2 + x^2} - a^2$$

(6)

where $a$ is a positive number.

This function was first proposed in [29] as a potential function where $a$ was fixed to 1. We give the scale parameter $a$ to it. This scale parameter is used to specify the error value at which the influence switches from behaving linearly to behaving like constant. The function $\rho(x, a)$ is successfully used in restoring the images from those corrupted by mixed impulse and Gaussian noise in [31]. The influence function $\psi(x, a)$ is the first derivative of $\rho(x, a)$ with respect to $x$

$$\psi(x, a) = \frac{a \sqrt{a^2 + x^2}}{a^2 + x^2}$$

(7)

As mentioned above, this function characterizes the bias that a particular error has on the solution.
With the parameter $a$ fixed, $\rho(x, a) \approx \frac{1}{2} x^2$ when $x$ is small enough and $\rho(x, a) \approx a|x| - a^2$ when $x$ is relatively large. The curves of $L_2$ norm, $L_1$ norm, $\rho(x, 1)$ and their corresponding influence functions are plotted in Fig. 2. The influence function $\psi(x)$ gradually approaches $L_1$ influence function for $x \geq 1$. However, it differs from $L_1$ influence function by discriminating a range of small error values like $L_2$ norm. We can control this transition from $L_1$ norm to $L_2$ norm by modifying the parameter $a$. In this sense, $\rho(x, a)$ is an adaptive robust norm. Fig. 3 shows several this norms and their influence functions corresponding to a set of different $a$ values. We can see that the adaptive error norm takes the shape of $L_1$ norm when $a$ tends to zero. With a increasing, the adaptive error norm accepts a larger range of errors as the effect of noise and handles them like $L_2$ norm.

Using the proposed adaptive norm in Eq. (6), the fidelity term of our robust SR estimation is

$$F(X) = \sum_{k=1}^{p} \sum_{m=1}^{M} \rho(e_{k,m}, a_k)$$
$$= \sum_{k=1}^{p} \sum_{m=1}^{M} \left( a_k \sqrt{a_k^2 + e_{k,m}^2} - a_k^2 \right)$$

where $a_k$ is the threshold parameter for the $k$-th frame. The using of different thresholds for different LR frames is motivated by the different accuracy levels of each estimated LR image observation model. The same idea was used in [21,19] when considering different registration accuracies. Subtracting a constant has no influence to the minimization, we can obtain our solution by the following minimization problem

$$\hat{X} = \arg \min_X \sum_{k=1}^{p} \sum_{m=1}^{M} a_k \sqrt{a_k^2 + e_{k,m}^2}$$

(8)
3.3. Adaptive calculation of $a_k$

The precise registration of the subpixel motion and knowledge of the point spread function (PSF) are very important to the reconstruction of HR image. However, the accurate estimation of these parameters is a difficult task in real application. As a result, the residual errors of each LR frame inevitably suffer from three types of noise: blur noise due to inaccurate estimation of the PSF, registration noise due to inaccurate registration and additive Gaussian noise [3]. These three types of noise can result in different residual error levels per frame. For example, registration might be more successful for some LR frames than for others. Furthermore, the system PSF can be different between frames due to time-varying atmospheric turbulence. To address this problem in our proposed SR algorithm, we developed an adaptive strategy to determine $a_k$ for each LR frame according to its reliability. In this sense, the threshold $a_0$ can be seen as the indicator of the usefulness and importance of each frame. Frames with larger residual errors should be seen as having less contributions to the reconstructed HR image and be given smaller thresholds.

In this paper, the initial SR estimation $\hat{X}_0$ is obtained by bilinear interpolation of the first reference LR frame. Due to the different accuracy levels of the estimated observation models among LR frames, the averaged residual error $E_k = \|DB_k M_k \hat{X}_0 - Y_k\|_2^2$ for each frame $k$ may be different. When a certain LR frame contains no outliers and is precisely registered and PSF estimated, the corresponding residual error will be small and approximately obey Gaussian distribution. In this case, $L_2$ norm is still a good selection [3]. Therefore, we should set large $a_k$ values for such frames to handle them like $L_2$ norm. However, for those frames containing outliers or suffering from inaccurate subpixel registration and PSF estimation, the residual errors will be large and dominated by registration noise and blur noise. To discount such frames and suppress their contribution to the final HR solution, we should set small $a_k$ values to them. Note that $L_2$ norm can also be used in $E_k$ to measure the residual error. However, $L_2$ norm is more sensitive to large errors than $L_1$ norm, the existence of outliers or misregistration in some LR frame may bias the corresponding $E_k$ to be too large. This is the reason we prefer $L_1$ norm to $L_2$ norm in $E_k$.

Based on above analysis, we seek to incorporate the following desirable properties in our calculation of $a_k$: (i) $a_k$ is inversely proportional to $E_k$, (ii) $a_k$ is larger than zero, and if $E_k \to E_{\text{min}}$ (the minimum of all $E_k$ values), $a_k \to a_{\text{max}}$ (upper bound on $a_k$), (iii) if $E_k \to E_{\text{max}}$ (the maximum of all $E_k$ values), $a_k \to a_{\text{min}}$ (lower bound on $a_k$). Under these constraints, we define the function $a_k = \theta(E_k)$ where $\theta(\cdot)$ is a monotonically decreasing function.

There are many types of monotonically decreasing functions which have the above-mentioned properties. We consider the following quadratic function as a reasonable choice to calculate $a_k$ from $E_k$:

$$a_k = -\tau E_k^2 + \gamma$$

where parameter $\tau > 0$ controls the decay of this quadratic function. Given the last two constraints, parameters $\tau$ and $\gamma$ are calculated respectively as:

$$\tau = \frac{a_{\text{max}} - a_{\text{min}}}{E_{\text{max}}^2 - E_{\text{min}}^2}, \quad \gamma = \frac{a_{\text{max}} E_{\text{max}}^2 - a_{\text{min}} E_{\text{min}}^2}{E_{\text{max}}^2 - E_{\text{min}}^2}$$

The lower bound $a_{\text{min}}$ is chosen to be a small number. In our experiments presented in this paper, $a_{\text{min}}$ was set to 0.1. We set $a_{\text{max}}$ to be the maximal averaged residual error among all LR frames, that is $a_{\text{max}} = E_{\text{max}}$.

The reason for choosing $\theta$ to be the quadratic function lies in its linear derivative. The change rate of the quadratic function becomes larger as the residual error is larger. The larger residual errors correspond to those frames are badly registered and PSF estimated, and hence, the quadratic function is very sensitively responsive to these frames and set small thresholds to them. On the other hand, Since the initial estimation $\hat{X}_0$ is only an approximation to the original HR image, the residual errors also contain some novel information. The quadratic function has a smaller change rate when the residual error is small. This nature makes it not sensitively respond to such small residual errors containing novel information.

4. Proposed super-resolution algorithm

Generally, the SR image reconstruction is often an ill-posed inverse problem because of an insufficient number of LR images and ill-conditioned blur matrices [18]. Therefore, regularization technique is necessarily applied in SR to well pose this problem. The solution for the regularized SR methods is given by:

$$\hat{X} = \arg \min_X \{ F(X) + \lambda R(X) \} \quad (9)$$

where $F(X)$, called the fidelity term, measures the closeness of an estimated HR image to the captured LR images and has been discussed in above section. The term $R(X)$, called the regularization term, is utilized to regularize the problem and to achieve a stable solution to the problem. The scalar $\lambda$ is the regularization parameter to balance the weight between the fidelity term and the regularization term.

4.1. Bilateral edge-preserving regularization

Some of the widely used regularization functions for SR problem are Tikhonov-type regularizer [2–4,8,11,22–25]:

$$R_{\text{Tik}}(X) = \| \Gamma X \|_2^2$$

$$R_{\text{Tik,TV}}(X) = P_{\text{LOD}}(\Gamma X, T_l) \quad (10)$$

and Total Variation (TV)-type regularizer [6,7,32]:

$$R_{\text{TV}}(X) = \| \nabla X \|_1$$

$$R_{\text{BTV}}(X) = \sum_{l=-q}^{q} \sum_{m=-m}^{m} \alpha ||l+m||_1 \| X - S_x^l S_y^m X \|_1 \quad (11)$$

where $\Gamma$ is a highpass operator such as Laplacian, $\nabla$ is a gradient operator, $S_x^l$ and $S_y^m$ shift $X$ by $l$, and $m$ pixels in horizontal and vertical directions, respectively, and the scalar weight $\alpha$, $0 < \alpha < 1$, is applied to give a spatially decaying effect to the summation of the terms in BTV regularization. Other robust estimators including Gaussian, Leclerc, Hampel, Andrew’s Sine and Tukey’s Biweight estimators can also be used in Eq. (10) by replacing the Lorentzian estimator [22–24]. Due to their same properties summarized in Section 3.1, it is expected that they will have similar performance in regularized SR problem.

Edges are typically the most important features in an image. $R_{\text{Tik}}$ regularizer tries to limit the high-frequency component of the image and the resulting HR image will not contain sharp edges. According to the second property of the robust estimators, $R_{\text{Tik,TV}}$ regularizer does not penalize the high-frequency component larger than the setting threshold. Therefore, the reconstruction of the sharp edges in HR image sensitively depends on the initial setting at the start of the iteration. $R_{\text{TV}}$ and $R_{\text{BTV}}$ regularizers tend to preserve edges in the reconstruction, as they penalize gradient magnitudes measured by $L_1$ norm. Furthermore, BTV considers a
larger neighborhood and a spatially decaying effect when computing the gradient at certain pixel.

However, it is well known that regularization of the $L_1$ norm on gradient often leads to the piecewise constant result, and hence will produce artificial edges on the smooth areas. Charbonnier et al. developed an edge-preserving regularization scheme based on a class of potential functions [29]. Using the potential function to penalize image gradients, this type regularization can preserve large gradients corresponding to edges, while smooth small gradients which are often the effects of noise. Note that the function $\rho$ in Eq. (6) satisfies the conditions for edge preservation potential functions in [29]. We utilize this function to propose a new regularization term, called bilateral edge-preserving (BEP) regularization, looks like

$$ R_{\text{BEP}}(X) = \sum_{l=-q}^{q} \sum_{m=0}^{N} \alpha |l| + |m| \rho((X - S_1^y S_m^x X)[l], c) $$

\[(12)\]

with $\alpha$, $q$, $S_1^y$, $S_m^x$ the same in Eq. (11). Here $(X - S_1^y S_m^x X)[l]$ is the $i$-th element of the vector $X - S_1^y S_m^x X$, parameter $c$ is the threshold at which the penalty for gradient switches from severe to slight.

The proposed regularization term $R_{\text{BEP}}(X)$ uses edge-preserving potential function $\rho(x, c)$ to penalize the gradient, and hence can preserve sharp edges like TV regularization. Its advantage over TV regularizer is, by smoothing a range of small gradients like Tikhonov regularization, it can reduce artificial edges (discontinuities) on the continuous areas produced by TV regularization. We can control the transition from behaving like TV regularization to Tikhonov regularization by changing parameter $c$. Furthermore, $\rho(x, c)$ is strictly convex and continuously differentiable, so it does not involve numerical difficulties as when using TV and BTV regularization.

4.2. Proposed minimization model

This subsection describes the minimization model in our SR algorithm. The model differs from the others, mainly, by the utilization of the proposed adaptive norm both in the fidelity and regularization terms. The functional to be minimized is defined by

$$ C(X) = \sum_{k=1}^{p} \sum_{m=1}^{M} \alpha_k \sqrt{\alpha_k^2 + \varepsilon_k^2} + \lambda \sum_{l=-q}^{q} \sum_{m=0}^{N} \alpha |l| + |m| \rho((X - S_1^y S_m^x X)[l], c) $$

\[(13)\]

Steepest descent (SD) algorithm was used to solve the regularized SR problem in [8,11,22–24]. Although our minimization model (13) can also be solved by SD algorithm, a conjugate gradient (CG) optimization algorithm usually converges much faster [33,30]. Moreover, the strict convexity of the function in Eq. (6) implies the strict convexity of the entire functional in Eq. (13), and therefore guarantees the CG optimization technique converge to the global minimizer. The procedure for CG optimization of our model (13) is described as follows. The current HR estimate is updated by

$$ \hat{X}_{n+1} = \hat{X}_n + \alpha_n P_n, \quad n = 1, 2, \ldots $$

where $P_n$ is the conjugate-gradient vector at the $n$-th iteration. By initializing the conjugate-gradient vector as $P_0 = -\nabla C(\hat{X}_0)$, all subsequent vectors are computed by

$$ P_{n+1} = -\nabla C(\hat{X}_{n+1}) + \beta_{n+1} P_n, \quad \beta_{n+1} = \frac{\nabla C(\hat{X}_{n+1})^T \nabla C(\hat{X}_{n+1})}{\nabla C(\hat{X}_n)^T \nabla C(\hat{X}_n)} $$

5. Experiments

In this section, we use both artificially generated and real data to test the proposed algorithm, and compare its performance with some other existing approaches. The generated data allow us to justify the different parts of our algorithm in Eq. (13) separately. The peak signal-to-noise ratio (PSNR) is used to measure the quality of the estimated HR image. This metric is defined as PSNR = $10 \log_{10}$[[255]2/\|\hat{X} - X\|^2]], where $L_1 N_1 L_2 N_2$ is the total number of pixels in the HR image, and $\hat{X}$ and $X$ are the reconstructed HR image and the original image, respectively. We then employ a real data set to compare our SR algorithm with other popular algorithms. All of the iterative initial values are obtained by bilinear interpolation of the first reference LR frames. We select the regularization parameters to produce both most visually appealing results and the highest PSNR in all the experiments below. Each PSNR value reported in all tables in this section is the best result of several runs with different regularization parameters.

5.1. Simulation experiments

We start with two synthetic sequences (pepper and cameraman) with global motion to justify our proposed fidelity and regulariza-
larization terms in Eq. (13) separately. Then we compare our proposed SR algorithm with other methods in the literature by using the same data. Fig. 4 shows the original pepper and cameraman image with size of $256 \times 256$ and $128 \times 128$, respectively. Sixteen LR pepper images are acquired by shifting Fig. 4(a) in vertical and horizontal directions, blurring it with a $5 \times 5$ Gaussian blur kernel with standard deviation equal to 1.5, and subsampling it by factor of four. The six LR cameraman images are generated in the same way, but the PSF of the blur is a $3 \times 3$ Gaussian kernel with standard deviation equal to 1 and the downsampling factor is 2.

5.1.1. Effectiveness of the BEP regularizer

Firstly, by fixing the fidelity term, we justify the effectiveness of the proposed BEP regularizer and compare it with the Tikhonov, BTV and the Lorentzian–Tikhonov regularizers. There are several types of $\rho$–Tikhonov regularizers in [11,22–25], where $\rho$ is the robust M-estimators including Gaussian, Lorentzian, Hample, Leclerc,
5.1.2. Effectiveness of the adaptive fidelity term

We further justify the effectiveness of the proposed adaptive fidelity term by comparing it with the $L_2$, $L_1$, Huber and Leclerc fidelity terms. Note that we take the Leclerc estimator as an example of the robust estimators this time. We compare the performance of all the fidelity terms without regularization. Firstly, to study their robustness against outliers, we single out one LR pepper image and contaminate it with salt and pepper noise at three levels (1%, 5% and 10%). The PSNR values of all the estimated HR images are shown in Table 2. Fig. 6 shows the zoomed HR images when the salt and pepper noise level is 10%. As can be seen from the table and the figure, $L_2$ fidelity term is not robust against outliers. The Leclerc fidelity term with $T = 1$ has the best performance in rejecting outliers. This is because Leclerc’s influence function tends to zero when the error is larger than the fixed threshold $T = 1$. From this point of view, the performance of the Leclerc fidelity term depends on the choice of the threshold. When its threshold $T$ is set to be 4, Leclerc fidelity term is outperformed by our proposed adaptive fidelity term.

Secondly, we test the performance of all the fidelity terms in handling registration errors and PSF estimation errors. To simulate the errors due to misregistration, we single out one LR pepper image and introduce a deliberate error in its registration parameter corresponding to a translation error of 4 pixels on the HR image grid. The PSF is assumed to be a normalized $5 \times 5$ Gaussian kernel of zero mean and unity variance to simulate blur estimation errors. The resulting PSNR values and zoomed HR images are shown...
Fig. 7. Results obtained by applying various fidelity terms to LR pepper images (including registration errors and PSF estimation errors). (a) One LR image. (b) Iterative initial (result of bilinear interpolation). (c) Result of $L_2$ fidelity. (d) Result of $L_1$ fidelity. (e) Result of Lorentzian fidelity ($T = 1$). (f) Result of Lorentzian fidelity ($T = 4$). (g) Result of Huber fidelity ($T = 1$). (h) Result of Huber fidelity ($T = 4$). (i) Result of proposed adaptive fidelity.

5.1.3. Effectiveness of the proposed SR algorithm

In this subsection, we compare our proposed SR algorithm with the methods in the literature. Firstly, we add the Gaussian noise with $\sigma = 5$ and $\sigma = 10$ to all the generated LR pepper and cameraman images. Note that we do not introduce the outliers, registration or PSF estimation errors in this experiment. In the second experiment, we introduce the outliers, registration and PSF errors and additive Gaussian noise to the SR problem. The outliers are simulated by adding the salt and pepper noise with two levels (5% and 10%) to one LR frame. We obey the same ways that we did in the last subsection to introduce the registration and PSF errors to both images. The Gaussian noise is added to all LR images to achieve signal-to-noise (SNR) ratio equal to 30 db. This metric is defined as $\text{SNR} = 10 \log_{10} \left( \frac{\| X \|^2}{N \sigma^2} \right)$, where $\sigma^2$ is the variance of the noise and $N$ is the size of the image $X$. The PSNR values of the estimated HR images by various methods are shown in Table 3.

We can conclude that, as we expected, the robust estimator based regularization methods in [11, 22, 24] have similar performance in solving SR problems. This is because all of the robust estimators have the same properties that we summarized in Section 3.1. Our proposed algorithm has the best performance in reconstruction in terms of the PSNR values. To see the visual effects of the estimated HR images, we show parts of them in Fig. 8. The HR result of the conventional $L_2 +$ Tikhonov regularization is shown in Fig. 8(c). As $L_2$ norm is not robust against outliers and large registration errors, it is not surprising to see the artifacts at the location of the salt and pepper noise and shadows around cameraman’s head. Figs. 8 (d) and (e) show the reconstructed HR images by the $L_1 +$ BTV algorithm [7] and the Lorentzian + Lorentzian–Tikhonov algorithm [11]. The reconstructed HR image by our proposed algorithm is shown in Fig. 8(f). As can be seen, our result can adequately suppress the negative effects of outliers and errors caused by misregistration and PSF misestimation, and is generally
Table 3
PSNR (dB) results obtained by applying various methods to both images.

<table>
<thead>
<tr>
<th>Method</th>
<th>Pepper</th>
<th>Cameraman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gaussian</td>
<td>Salt pepper</td>
</tr>
<tr>
<td></td>
<td>(\sigma = 5)</td>
<td>(\sigma = 10)</td>
</tr>
<tr>
<td>(L_1 + R_{Tik})</td>
<td>28.31</td>
<td>26.40</td>
</tr>
<tr>
<td>(L_1 + R_{BTV} [7])</td>
<td>28.63</td>
<td>27.07</td>
</tr>
<tr>
<td>(L_2 + R_{Tik})</td>
<td>29.04</td>
<td>27.58</td>
</tr>
<tr>
<td>(L_2 + R_{BTV})</td>
<td>30.29</td>
<td>28.90</td>
</tr>
<tr>
<td>Huber + R_{HamTik} [8]</td>
<td>30.27</td>
<td>28.92</td>
</tr>
<tr>
<td>Proposed algorithm</td>
<td>30.35</td>
<td>29.12</td>
</tr>
</tbody>
</table>

Fig. 8. Results obtained by applying various methods to LR cameraman images reconstruction (salt pepper noise level = 10\%). (a) One LR image corrupted with 10% salt and pepper noise and Gaussian noise. (b) Iterative initial. (c) Result of \(L_2 + R_{Tik}\). (d) Result of \(L_1 + R_{BTV}\) in [7]. (e) Result of Lorentzian + R_{LorTik} in [11]. (f) Result of proposed algorithm.

more visually appealing without destroying fine details in the image.

5.2. Practical experiment

We also use the proposed SR algorithm on a real “Alpaca” data sequence, which is composed of 55 (32 x 70) LR frames. This data sequence can be downloaded from Milanfar’s web site. This experiment differs from above simulation experiments in three ways. First, subpixel motion vectors between LR frames are not defined explicitly. Second, the motions are more complicated: the first 45 frames approximately follow the global translational motion model; the last 10 frames follow a more complicated motion model including Alpaca movement. Thirdly, the accurate camera PSF is unknown. We use the method described in [35] to estimate the motion vectors and the unknown camera PSF is assumed to be a 3 x 3 Gaussian kernel with standard deviation equal to 1. The resolution enhancement factor was set to be 3.

Figs. 9 (a) and (b) show the first LR frame and the 53-th LR frame (including Alpaca’s movement), respectively. \(L_2\) norm reconstruction with Tikhonov regularization result is shown in Fig. 9(c). From this result it is shown \(L_2\) norm estimate suffers from visible artifacts that are caused by Alpaca’s movement. Fig. 9(d) shows the HR estimate using the Lorentzian + Lorentzian–Tikhonov regularization algorithm proposed in [11]. From this result it is shown the use of the Lorentzian error norm has suppressed the visible artifacts. However, the solution is of relatively poor quality. The HR estimate obtained by Farsiu’s \(L_1 + BTV\) algorithm [7] is shown in Fig. 9(e). Fig. 9(f) shows the result using the proposed algorithm. From this result, it is observed that the proposed algorithm successfully suppress the influence of the Alpaca’s movement, result in the HR solution with crisper details than Farsiu’s. Plots of the averaged residual errors \(E_k\) and the outlier thresholds \(a_k\) are shown in Figs. 9 (g) and (h), respectively. From these two plots it can be observed that the estimated LR observation models and real models. It is also observed that the outlier thresholds corresponding to the last ten frames are significantly small because of their violation of the assumed translational motion model.

1 Available at http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html.
Fig. 9. HR estimates for the test “Alpaca” sequence. (a) The first LR frame. (b) The 53-th LR frame. (c) Result of $L_2$ fidelity term + Tikhonov regularization. (d) Result of Lorentzian fidelity term + Lorentzian–Tikhonov regularization [11]. (e) Result of $L_1$ fidelity term + BTV regularization in [7]. (f) Result the proposed algorithm. (g) Plot of the residual error values $E_k$. (h) Plot of the outlier thresholds $a_k$.

6. Conclusion

In this paper, a new algorithm based on regularization framework is proposed to address multiframe super-resolution image reconstruction problem. The proposed algorithm differs from the others by using an adaptive error norm both in the fidelity and regularization terms. Using this proposed error norm in the fidelity term, we developed an adaptive strategy to effectively deal with violations to the assumed imaging model. We also presented a new regularization term based on the proposed error norm. This new BEP regularization term can be seen as the improvement of the BTV regularizer. Experiments have been performed on both the artificially generated and real data, and the proposed algorithm has demonstrated better results in both cases, compared with other methods in the literature.

Acknowledgments

The authors would like to thank the anonymous reviewers for helping to improve this paper. This research was partially supported by National Natural Science Foundations of China (Nos. 11071261, 91130009), the “Computational Science Innovative Research Team” program and Guangdong Province Key Laboratory of Computational Science at the Sun Yat-sen University.

References