This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

Regional Spatially Adaptive Total Variation Super-Resolution with Spatial Information Filtering and Clustering

Qiangqiang Yuan, Liangpei Zhang, *Senior Member*, *IEEE,* Huanfeng Shen, *Member*, *IEEE*

Abstract—Total variation has been used as a popular and effective image prior model in the regularization-based image processing fields. However, as the total variation model favors a piecewise constant solution, the processing result under high noise intensity in the flat regions of the image is often poor, and some "pseudo-edges" are produced. In this paper, we develop a regional spatially adaptive total variation (RSATV) model. Firstly, the spatial information is extracted based on each pixel, and then two filtering processes are respectively added to suppress the effect of "pseudo-edges". After that, the spatial information weight is constructed and classified with k-means clustering, and the regularization strength in each region is controlled by the clustering center value. The experimental results, on both simulated and real datasets, show that the proposed approach can effectively reduce the "pseudo-edges" of the total variation regularization in the flat regions, and maintain the partial smoothness of the high-resolution image. More importantly, compared with the traditional pixel-based spatial information adaptive approach, the proposed region-based spatial information adaptive total variation model can better avoid the effect of noise on the spatial information extraction, and maintains robustness with changes in the noise intensity in the super-resolution process.

Index Terms—Super-resolution, total variation, regional spatially adaptive, majorization-minimization

I. INTRODUCTION

High-resolution (HR) imagery plays a key role in many diverse areas of application, such as medical imaging [1], remote sensing [2], [3], and video surveillance [4]. However, because there are a number of limitations with both the theoretical and practical aspects, such as the sensor resolution and high cost, amongst other things, it is obviously more difficult to obtain a HR image than a low-resolution (LR) image. Consequently, researchers have explored ways to produce a HR image from the image processing aspect, and, in recent decades, super-resolution (SR) technology, which produces a HR image from single-frame or multi-frame LR images, has been proposed. In this paper, our research is mainly focused on the multi-frame image SR problem: the process of reconstructing a HR image from a sequence of LR images.

A. Problem Formulation

Assume that a HR image x is shifted, blurred, downsampled, has some additive noise, and produces a sequence of LR images y_k (Fig. 1). The standard image degradation model for the problem of SR is in the form:

$$y_k = DB_kM_kx + n_k \quad k = 1,\cdots,p \tag{1}$$

Qiangqiang Yuan is with the school of Geodesy and Geomatics, Wuhan University, P.R. China(e-mail: yqiang86@gmail.com).

Liangpei Zhang is with are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, P.R. China. (e-mail: zlp62@whu.edu.cn).

Huanfeng Shen is with the School of Resource and Environmental Science, Wuhan University, P.R. China(e-mail: shenhf@whu.edu.cn).

This work was supported by Major State Basic Research Development Program (973 Program) of China under Grant 2011CB707103, the National Natural Science Foundation of China under Grants 40930532, 41271376, 61201342, the National Science and Technology Support Program under Grant 2012BAJ23B03, and the China Postdoctoral Science Foundation under Grant 2012M521471.

where y_k is the kth LR image rearranged in lexicographic order, which has the size of $L_1L_2 \times 1$, and x is the original HR image, which is also rearranged in lexicographic order, and which has the size of $H_1H_2 \times 1$. M_k stands for the warp matrix with size $H_1H_2 \times H_1H_2$, B_k is the blurring matrix with size $H_1H_2 \times H_1H_2$, D is the downsampling matrix with size $L_1L_2 \times H_1H_2$, vector n_k is the system noise with size $L_1L_2 \times 1$, which, in this paper, is assumed to be Gaussian additive noise with zero mean, and p represents the number of LR images. In this paper, we assume that the blurring matrix B_k remains the same between the LR images y_k. The purpose of multi-frame SR reconstruction is to reconstruct the HR image x from sequences of LR images y_k, as shown in Fig. 1.

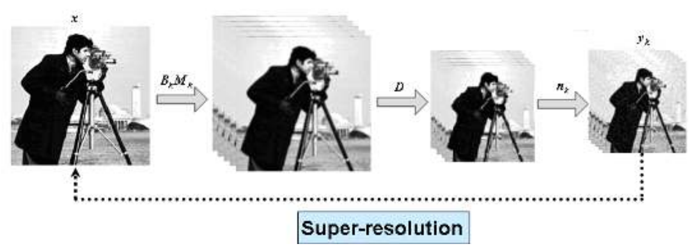


Fig. 1. The degradation process of the HR image, and the super-resolution process.

B. Previous Algorithms

In recent decades, the multi-frame SR problem has been widely explored by many researchers, and considerable progress has been achieved. Tsai and Huang [5] first proposed to use multi-frame SR theory to enhance the resolution of multi-temporal Landsat TM images in the frequency domain. After that, many other improved frequency domain SR algorithms have also been proposed [6], [7]; however, for the frequency domain approaches, although they have the advantage of a short computation time, it is difficult to add the prior information of the HR image. Therefore, researchers have attempted to solve the SR problem in the spatial domain, and various algorithms have been developed, such as the projection onto convex sets (POCS) approach [8], [9], maximum likelihood (ML) approach [10], maximum a posteriori (MAP) approach [11], [12], joint MAP approach [13]–[14], and the hybrid approach [15]. Recently, some excellent SR algorithms that do not rely on exact motion estimation have been proposed [16], [17], and very promising SR results were produced, especially when complex motions are contained in the LR image sequences. In addition, wavelet domain SR methods have also been proposed [18], [19]. Reviews of the state of the art of SR methods can be found in [20]–[24].

Because SR is an ill-posed problem, it is wise to incorporate some prior distribution of the HR image to constrain the SR process and obtain a stable and relative optimal solution. Therefore, in recent decades, many prior models of the HR image have been proposed. The most widely used prior model is the Tikhonov regularization model [13], which is used to guarantee the smoothness property of the original HR image. However, although the Tikhonov model is simple to realize and easy to solve, it has the drawback of blurring the edges Therefore, many edge-preserving prior image models have been proposed, including the Huber-Markov random field (Huber-MRF) model [11], total variation (TV) model [25]–[27], bilateral total variation (BTV) model [21], and the weighted Markov random field (WMRF) model [4]. Recently, sparse representation-based prior models have been proposed and have shown very promising single image restoration and SR results [28]–[31]. Among these models, the TV model is a very popular one because of its strong ability of edge preserving. However, the traditional TV model also has its shortcoming in that because it assumes that the image is piecewise smooth, some "pseudo-edges", which are also called the "staircase effect", may be produced in the smooth regions, especially in high noise or blur conditions [32].

Therefore, to overcome the shortcoming mentioned above, some spatially adaptive TV (SATV) models, which use the spatial information to constrain the regularization strength in each pixel, have been developed. The basic idea of the spatially adaptive regularization model is to use the spatial information distributed in the image to constrain the regularization strength. A weak regularization strength is enforced in the edge pixels to preserve detail information, and a strong regularization strength is enforced in the homogeneous area pixels to effectively suppress noise. The first spatially adaptive idea for a TV model can be attributed to Strong et al. [33], where the authors proposed to use a gradient image to constrain the TV regularization strength in different pixels. A weak regularization strength is enforced in the edge pixels with a large gradient to preserve detail information, and a strong regularization strength is enforced in the flat area pixels with a small gradient to effectively suppress noise and the "pseudo-edges". Clearly, the performance of this model is largely dependent on the gradient information extraction process. Because the gradient information is based on a pixel unit, if high-intensity noise is included in the observation image, a noise pixel will be falsely recognized as an edge pixel and a weak regularization strength will be enforced, which will cause the noise and "pseudo-edges" in the flat regions to be poorly suppressed. Recently, Chen et al. [34] proposed a new edge indicator called "difference curvature", instead of the gradient information, to further improve Strong's method. However, although the difference curvature indicator works better than the gradient information, it is also based on a pixel unit, and cannot work well in high noise intensity conditions. Guo et al. [35] proposed a local mutual information weighted TV model to denoise a magnetic resonance image (MRI), but this approach needs a regulating image to compute the local mutual information, which limits it to use with MRI images. Under the variational Bayesian framework, Chantas et al. [36] developed a product of a spatially weighted TV model, in which the image restoration and spatially weighted parameter estimation are executed simultaneously. Chopra et al. [37] proposed to adapt the smoothly clipped absolute deviation (SCAD) penalty theory from the statistical community to improve the SATV model. In addition, the SATV model has also been used on color image sharpening-demosaicking problems [38].

C. Proposed Algorithm

In this paper, we aim to construct the spatial constraint from a regional perspective, and a regional spatially adaptive total variation (RSATV) model is proposed. The main idea and contribution of the RSATV model can be concluded as follows, and the outline of the proposed RSATV model is specifically presented in Fig. 2.

To suppress the effect of the noise, a median filter process is enforced on the pixel-based spatial information, before it is used to construct the spatial weight, After the spatial weight is computed, it is mean filtered and classified with the k-means clustering method, and the spatial weight in different image regions is defined with the cluster center value of each spatial information class, instead of each pixel, as in the traditional SATV models. This means that for different regions, different regularization strengths are enforced, which maintains the homogeneous nature of the spatial information and further suppresses the effect of noise. For the optimization process, the majorization-minimization (MM) algorithm is adopted. In each iteration, both filtering and clustering processes are executed, and the spatial weight is updated iteratively, which maintains a more accurate and robust regional spatial information constraint.

D. Organization of the Paper

The remainder of this paper is organized as follows. The regularization-based SR model is described in Section II. In Section III, our RSATV model is presented in detail. The optimization process is described in Section IV. In Section V, the experimental results and a discussion are presented and, finally, conclusions are drawn in Section VI.

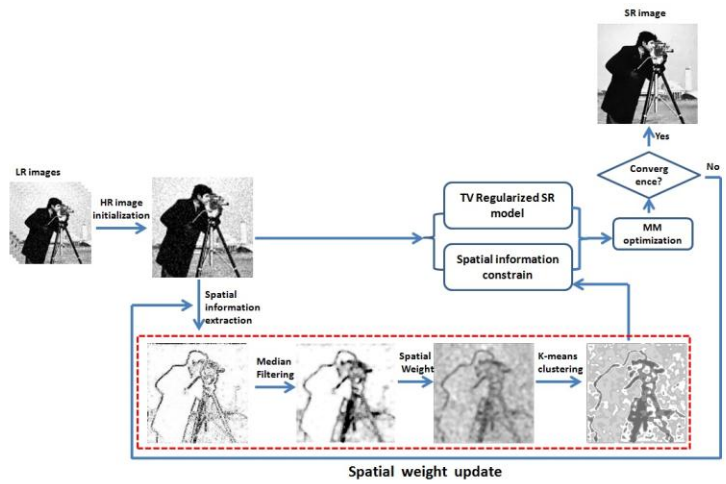


Fig. 2. Outline of the proposed RSATV model (the red dashed line presents our contribution).

II. TOTAL VARIATION REGULARIZED SUPER-RESOLUTION

In this section, we introduce the regularized SR model and the TV regularization model.

A. The Regularized Super-Resolution Model

For the degradation model presented in (1), because the SR process is an ill-posed problem, some prior information about the HR image should be added to guarantee a stable and relative optimal solution. A popular and effective approach to this problem is to use the regularization-based least squares method, which has the following formulation [39]:

$$\hat{x} = \arg\min_{x}\{\sum_{k=1}^{p}\left\| y_k - DBM_k x \right\|_2^2 + \lambda P(x)\} \qquad (2)$$

In (2), p represents the number of LR images, $\sum_{k=1}^{p}\left\| y_k - DBM_k x \right\|_2^2$ is the data fidelity item, which stands for the fidelity between the observed LR image and the original HR image, and $P(x)$ is the regularization item, which gives a prior model of the HR image x. λ is the regularization parameter, which controls the trade-off between the data fidelity and prior item.

B. Total Variation Regularization

In (2), the regularization item $P(x)$, which stands for the prior distribution of the HR image, plays a very important role in the SR process. It controls the perturbation of the solution, solves the ill-posed problem for SR reconstruction, and guarantees a stable HR estimation. Among the many proposed prior models, the TV model is very popular and effective because of the property of edge preserving [40]. For the HR image x, the TV model can be defined as follows:

$$TV(x)=\sum_i \sqrt{\left(\nabla_i^h x\right)^2+\left(\nabla_i^v x\right)^2} \quad (3)$$

where ∇_i^h and ∇_i^v are linear operators corresponding to the horizontal and vertical first-order differences, respectively. At pixel i, $\nabla_i^h x = x_i - x_{r(i)}$, $\nabla_i^v x = x_i - x_{b(i)}$, and $r(i)$ and $b(i)$ represent the nearest neighbor to the right and below the pixel, respectively. The TV model presented in (3) is often defined as an isotropic total variation model, which means it is unaffected by rotation, reflection and changes in the position of an image [41]. Correspondingly, the anisotropic TV model is also defined as:

$$TV(x)=\sum_i \left|\nabla_i^h x\right|+\left|\nabla_i^v x\right| \quad (4)$$

Usually, an isotropic TV model is preferred over the anisotropic ones [41], [42]. Therefore, in this paper, an isotropic TV model is used.

Substituting $P(x)$ in (2) by $TV(x)$, as presented in (3), the TV SR problem can be written as:

$$\hat{x}=\arg\min_x\{\sum_{k=1}^{p}\left\|y_k-DBM_k x\right\|_2^2+\lambda TV(x)\} \quad (5)$$

III. RSATV: Regional Spatially Adaptive Total Variation Model

In this section, our regional spatially adaptive total variation (RSATV) model is introduced in detail.

A. *Spatial Information Extraction and Filtering*

For a given image x, we first extract the edge information distributed in the image. In this paper, the difference curvature indicator proposed in [34] is used. It has been proved that this indicator can effectively distinguish edges from flat and ramp areas in the image, and it performs better than the traditional gradient operator [34]. The definition of the difference curvature indicator is introduced as follows:

For the ith pixel in the image x, the difference curvature C_i is defined as:

$$C_i=\left\|u_{\eta\eta}\right|-\left|u_{\varepsilon\varepsilon}\right\| \quad (6)$$

where:

$$u_{\eta\eta}=\frac{u_x{}^2u_{xx}+2u_xu_yu_{xy}+u_y{}^2u_{yy}}{u_x{}^2+u_y{}^2} \quad (7)$$

$$u_{\varepsilon\varepsilon}=\frac{u_y{}^2u_{xx}-2u_xu_yu_{xy}+u_x{}^2u_{yy}}{u_x{}^2+u_y{}^2} \quad (8)$$

where η and ε are the direction of the gradient and the direction perpendicular to the gradient. In (7) and (8), u_x, u_y, u_{xx}, u_{yy} and u_{xy} stand for the first and second derivative gradient information of the pixel, respectively. $|\ |$ denotes the absolute value operator. $u_{\eta\eta}$ and $u_{\varepsilon\varepsilon}$ represent the second derivatives in the direction of the gradient ∇u and in the direction perpendicular to ∇u. The behavioral analysis of the difference curvature can be concluded as follows [34]:

1) For edges, $|u_{\eta\eta}|$ is large and $|u_{\varepsilon\varepsilon}|$ is small, so C_i is large.

2) For flat and ramp regions, $|u_{\eta\eta}|$ and $|u_{\varepsilon\varepsilon}|$ are both small, so C_i is small.

3) For noise pixels, $|u_{\eta\eta}|$ and $|u_{\varepsilon\varepsilon}|$ are both large, so C_i is small.

Before using the difference curvature information to construct the spatial weight, the difference curvature information is filtered with a median filter. The reason why the median filter is used will be explained specifically in the last part of this section. For example, for a 3×3 neighborhood window around $C_{i,j}$ (Fig. 3), the median value of its neighborhood pixel is selected as the filtering result.

$$V_i=median(C_{i-1,j-1},C_{i-1,j}\cdots C_{i+1,j},C_{i+1,j+1}) \quad (9)$$

$C_{i-1,j-1}$	$C_{i-1,j}$	$C_{i-1,j+1}$
$C_{i,j-1}$	$C_{i,j}$	$C_{i,j+1}$
$C_{i+1,j-1}$	$C_{i+1,j}$	$C_{i+1,j+1}$

Fig. 3. The neighborhood of $C_{i,j}$

B. *Spatially Weighted Parameter Construction*

After the spatial information is extracted and filtered, the following process is used to relate the spatial information of each pixel to the regularization strength of the TV model. In this paper, we construct a spatially weighted parameter W_i for each pixel, as follows:

$$W_i=\frac{1}{1+\beta V_i} \quad (10)$$

where V_i is the median filtered difference curvature value of pixel i, and β is a contrast factor. For the TV model in (3), we add the spatially weighted parameter in the following way:

$$PSATV(x)=\sum_i\left(W_i\times\sqrt{\left(\nabla_i^h x\right)^2+\left(\nabla_i^v x\right)^2}\right) \quad (11)$$

$$\hat{x}=\arg\min_x\{\sum_{k=1}^{p}\left\|y_k-DBM_k x\right\|_2^2+\lambda\times PSATV(x)\} \quad (12)$$

From (10), (11) and (12), it is shown that the spatially weighted parameter W_i can adaptively adjust the regularization strength of the TV model in different pixels in the image. For the flat region pixels, because the difference curvature V_i is small, the weighted parameter W_i will be large, and a strong regularization strength will be enforced on them to suppress noise. Conversely, for edge region pixels, as the difference curvature V_i is large, the weighted parameter W_i will be small, and a weak regularization strength will be enforced on them to preserve edge information.

C. *Spatially Weighted Parameter Filtering and Clustering*

Next, to automatically extract the flat regions and realize the spatially adaptive idea from a regional perspective, we propose to classify W_i with a k-means clustering approach, which has also been used in clustering based denoising problems [42], [43]. We expect the pixel-based spatial weight W_i to be divided into not necessarily contiguous regions, and each region will contain pixels with a similar spatial weight. Finally, the regularization strength of each pixel in the same clusters is controlled by the cluster center value. In the following paragraphs, the detail procedure is given.

Firstly, the spatially weighted parameter W_i is smoothed with a mean filter to help the clustering:

$$U_i=\frac{1}{w^2}\sum_{i=1}^{w^2}W_i \quad (13)$$

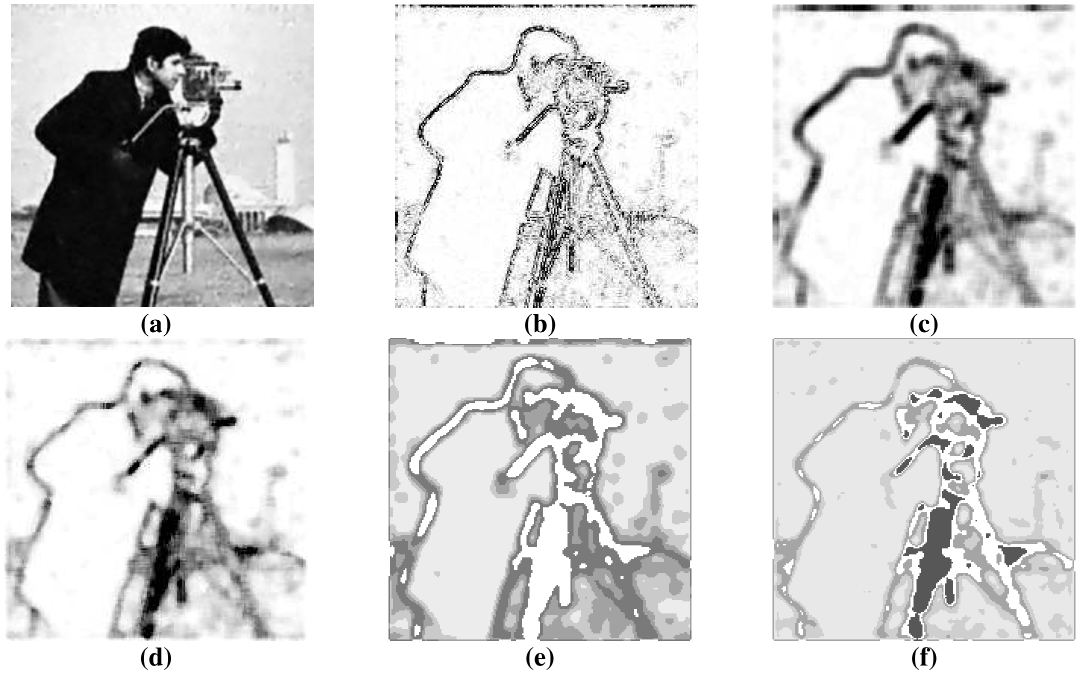


Fig. 5. **A comparison between a mean filter and a median filter in the iteration. (a) One of the HR images in the iteration; (b) the spatial information in (a); (c) the mean filter result on (b); (d) the median filter result on (b); (e) the regional spatial weight constructed from (c); and (f) the regional spatial weight constructed from (d).**

where w is the small window size used in the mean filter. After filtering, it is classified with k-means clustering.

It is assumed that, with k-means clustering, the spatially weighted parameter U_i is classified into n clusters, and the cluster center value of the cluster j is denoted as μ_j, $j=(1,2,\cdots n)$. For example, as is shown in Fig. 4, with the k-means clustering, the spatial weight U_i is classified into five clusters. The cluster with the largest cluster center value is determined to be a flat part, because using the k-means clustering approach, the pixels with a large spatial weight U_i will be grouped together, and the cluster center value will also be the largest. Therefore, for each cluster, the cluster center value μ_j is known and can be used to control the regularization strength. It can be clearly seen from Fig. 4 that, with the clustering process, a flat part can be automatically extracted and a large regularization strength can be easily enforced on it.

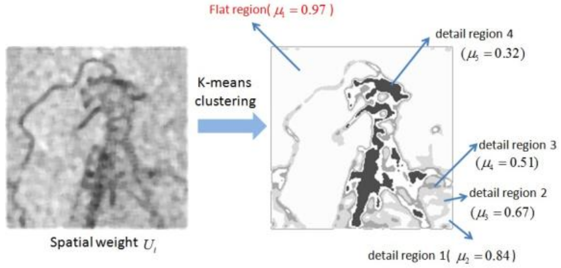


Fig. 4. The k-means clustering process of the spatial weight U_i.

For each region, because the cluster center value μ_j is known, the final spatial weight is defined as follows:

$$U_i^{region} = \begin{cases} \tau \times \mu_j & if \ \ i \in \textbf{flat region} \\ \mu_j & if \ \ i \in \ \textbf{detail region} \end{cases} \tag{14}$$

where τ is a constant parameter which helps to control the regularization strength and guarantee that a large regularization strength is enforced in the flat regions of the image.

From the above description, it can be seen that in the flat regions of the image, a large spatial weight value can enforce a strong regularization strength to suppress noise and the "pseudo-edges" phenomenon. Conversely, in the detail regions of the image, a small spatial weight parameter can guarantee a weak regularization strength to preserve the detail information.

After the spatially weighted parameter is constructed for each spatial region of the image, the RSATV model used in this paper can be defined as:

$$RSATV(x) = \sum_i \left(U_i^{region} \times \sqrt{\left(\Delta_i^h x\right)^2 + \left(\Delta_i^v x\right)^2} \right) \tag{15}$$

D. The Reason Why a Median Filter is Used

In Section III-A, a median filter is used to reduce the noise or artifacts in the spatial information C_i . The reason why a median filter is used here is strongly related to the structure of C_i . In our method, the HR image is solved iteratively, the spatial information is extracted in all the iterations , and the "pseudo-edges" are also reduced iteratively. Over the first few iterations, because of the effect of the noise, the spatial constraint is not accurate, causing the "pseudo-edges" to still exist in the reconstruction image. Consequently, the "pseudo-edges" will also be present in the spatial information C_i extracted from the SR image. For the "pseudo-edges", they can be more suitably reduced with a median filter than a mean filter smoothing. With a median filter.

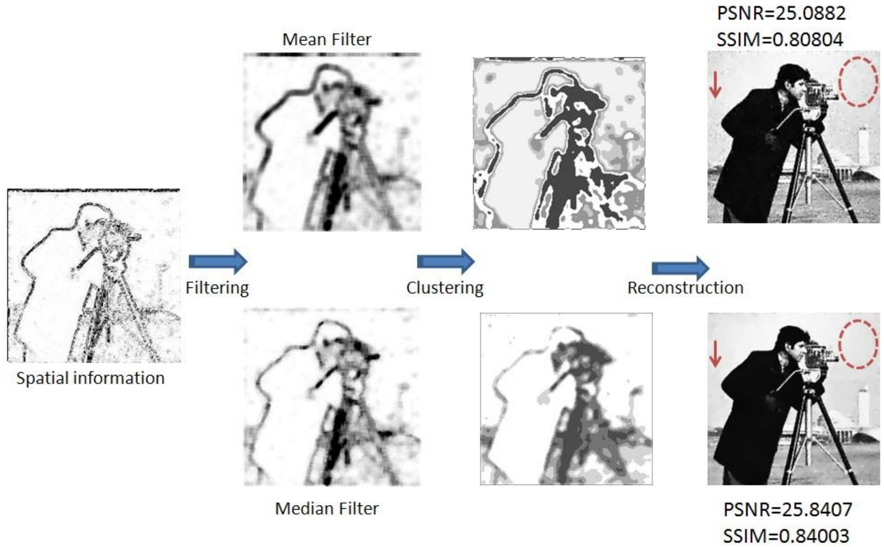


Fig. 6 A comparison between a mean filter and a median filter on the final SR results (noise variance = 18).

Fig. 5 is a comparison between a mean filter and a median filter in the iteration. From the figure, it can be seen that some "pseudo-edges" appear in the reconstruction image in each iteration, which can be seen more clearly from the spatial information extraction result C_i in (b). If a mean filter is used on C_i, we can see that although the "pseudo-edges" in the flat region are partially suppressed, the edge and texture information is smoothed. However, with a median filter, not only are the "pseudo-edges" in the flat region well suppressed, but this is also done without smoothing of the detail information. The advantage of a median filter is also reflected in the clustering results of (e) and (f), where it can be seen that the regional spatial weight produced by a median filter is more reasonable.

Fig. 6 is a comparison between the final reconstruction results with mean and median filters. This shows that a more accurate spatial constraint is produced with a median filter, and a better reconstruction image is also produced.

V. Majorization-Minimization Optimization

In this paper, the RSATV SR model is optimized with the MM approach proposed in [44]–[47]. The main idea of the MM optimization approach is to replace the traditional non-quadratic function with a quadratic and differentiable upper bound (majorization) equation, and then the optimization of the non-quadratic function can be replaced with the iterative optimization of the majorization equation [47].

To accomplish the MM idea with the RSATV model, we first consider the following relationship:

$$\sqrt{ab} \le \frac{a+b}{2} \rightarrow \sqrt{a} \le \frac{a+b}{2\sqrt{b}} \tag{16}$$

Let x^m be the current iterated image, and x is the HR image to be solved in the next iteration. Let $b = \left(\nabla_i^h x^m\right)^2 + \left(\nabla_i^v x^m\right)^2$, and $a = \left(\nabla_i^h x\right)^2 + \left(\nabla_i^v x\right)^2$.

Applying (16) to the RSATV model in (15), the functional majorization of the RSATV model in (15) can be defined as:

$$G_{RSATV}(x \mid x^m) = \left\{ \frac{1}{2} \sum_i \left\{ U_i^{region} \times \frac{\left(\nabla_i^h x\right)^2 + \left(\nabla_i^v x\right)^2 + \left(\nabla_i^h x^m\right)^2 + \left(\nabla_i^v x^m\right)^2}{\sqrt{\left(\nabla_i^h x^m\right)^2 + \left(\nabla_i^v x^m\right)^2}} \right\} \right\} \tag{17}$$

In (17), because x^m is known, (17) can be further written as:

$$G_{RSATV}(x \mid x^m) = \left\{ \frac{1}{2} \sum_i \left\{ U_i^{region} \times \frac{\left(\nabla_i^h x\right)^2 + \left(\nabla_i^v x\right)^2}{\sqrt{\left(\nabla_i^h x^m\right)^2 + \left(\nabla_i^v x^m\right)^2}} \right\} + C \right\} \tag{18}$$

Where C is a constant. Define θ_i^m, which has the formation:

$$\theta_i^m = \frac{1}{\sqrt{\left(\nabla_i^h x^m\right)^2 + \left(\nabla_i^v x^m\right)^2}} \tag{19}$$

Let $R = \begin{bmatrix} R^h \\ R^v \end{bmatrix}$, $Q^m = \begin{bmatrix} \Lambda^m & \\ & \Lambda^m \end{bmatrix}$ and $\Lambda^m = diag(\theta_i^m)$. R^h and R^v represent two matrices that have a size of $H_1H_2 \times H_1H_2$, such that $R^h x$ and $R^v x$ are the first-order differences of x. Equation (18) can be further written as:

$$G_{RSATV}(x \mid x^m) = \left\{ \frac{1}{2} \sum_i \left\{ U_i^{region} \times \theta_i^m \times \left(\left(R^h x\right)_i^2 + \left(R^v x\right)_i^2 \right) \right\} + C \right\}$$

(20)

Finally, (20) can be written as:

$$G_{RSATV}(x \mid x^m) = \left\{ (Rx)^T Q^m U (Rx) \right\} \tag{21}$$

$$U = diag(U_i^{region}) \tag{22}$$

Incorporating (21) into (2) and replacing $P(x)$ with $G_{RSATV}(x \mid x^m)$, the final functional majorization of the whole cost function can be written as:

$$G(x \mid x^m) = \{ \sum_{k=1}^{p} \left\| y_k - DBM_k x \right\|_2^2 + \lambda \times ((Rx)^T Q^m U (Rx)\} \tag{23}$$

For (23), because it is quadratic and differentiable, minimization with respect to x leads to the following linear system:

$$(\sum_{k=1}^{p}(DBM_k)^T(DBM_k)+\lambda\times((R)^TQ^mU(R))x^{(m+1)} = \sum_{k=1}^{p}(DBM_k)^Ty_k \quad (24)$$

For (24), the conjugate gradient (CG) algorithm can be used for the optimization. In each MM iteration, the regional spatially weighted parameter U_i^{region} is updated. It is also important to mention that the MM framework for the RSATV model will lead to the same result as the iterative reweighted norm (IRN) algorithm proposed in [48], which can also be used to optimize the RSATV model.

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experiment Setting

① Experiment Data

In our experiments, four simulated datasets and two real experiment datasets are used to verify the effectiveness of the proposed algorithm. The dynamic range of the six datasets is between 0 and 255.

The four original HR images used are, respectively, the "cameraman" image, with a size of 200×200, the "aerial" image, with a size of 200×200, the "Barbara" image, with a size of 256×256, and the "house" image, with a size of 256×256. The four original HR images are respectively shown in Fig. 7 (a)–(d). In the real data experiments, two datasets are used to verify the proposed algorithm. One dataset is the "EIA" image sequence obtained from the Multidimensional Signal Processing (MDSP) Research Group of UCSC [49], which consists of 16 frames with a size of 90×90. The other dataset is the "surveillance" video sequence, which was also provided by the MDSP Research Group of UCSC, and consists of 15 frames with a size of 66×76. In order to reduce the computational load, we just select the first 10 frames in the two real datasets. The well-performing registration approach presented in [50] is used as the motion estimation method.

(a) (b) (c) (d)

Fig. 7. The original images used in the simulated experiments: (a) the "cameraman" image, (b) the "aerial" image, (c) the "Barbara" image, and (d) the "house" image.

② SR Quality Evaluation Index

In the simulated experiments, we use the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) index to evaluate the simulated reconstruction results. The PSNR is employed to evaluate the gray value similarity, and the SSIM index, as proposed by Wang et al. [51], [52], is used to evaluate the structural similarity.

③ Parameter Setting

For the proposed RSATV model, in all the experiments, the filtering window size in the median filter process in (9) is set at 7×7, and the mean filter window size in (13) is set at 3×3. The regularization parameter λ and the parameter τ in (14) are adjusted until the best SR results are obtained. In all the experiments, the parameter β in (10) is set at 0.05, and the spatial weight parameter cluster number n (see Section III-C) is set at 5. In Section V-C, we also present a discussion on the effects of the parameter τ and cluster number n on the final SR performance, and give some advice about the setting of these parameters.

The termination condition of the CG procedure is set at 1e-5, and the termination condition of the MM procedure is also set at 1e-5. The resolution enhancement factor is set at 2 in all the experiments

B.Experimental Results

①Simulated Image Denoising Experiments

For the degradation model in (1), if we do not consider the motion, blurring, and downsampling processes, and just consider the additive noise, the model will become:

$$y = x + n \quad (25)$$

Then the SR problem is degraded to an image denoising problem, and the regularization-based denoising model can be expressed as follows:

$$\hat{x} = \arg\min_x\{\|y-x\|_2^2+\lambda P(x)\} \quad (26)$$

Therefore, the proposed RSATV model is first tested on an image denoising problem. To verify its effectiveness, it is compared with the TV model in [26], and some other spatially adaptive models, including the SATV model in [33], the SCAD model in [37], the ATV model in [34], and the PSATV model in equation (11). The "aerial" image is selected as the experiment data. In all these methods, the regularization parameter is adjusted until the best denoising result is obtained.

To test the noise robustness of the different algorithms, the denoising results with different noise variances (8, 14 and 18, respectively) are given. In Fig. 8, the denoising results with a noise variance of 18 are shown, and in Table II, the quantitative evaluation results under all the different noise conditions are given. From Fig. 8, it can be seen the proposed RSATV model gives the best denoising results among the five spatially adaptive TV models. In the other four models, because the spatial information is all extracted with a pixel unit, and the extraction process is deeply affected by noise pixels, which results in the spatial constraint being uncorrected, the noise in the flat regions is not well suppressed. In the high noise intensity condition, in particular, the pixel-based spatially adaptive TV model performs even worse than the traditional TV model. However, in the RSATV model, because the spatial information is filtered, and, meanwhile, the spatial constraint is enforced with a region unit with the help of the k-means clustering process, the noise in the flat regions is well reduced, and the edge and texture information is also well preserved. The better performance of the RSATV model can also be seen in the quantitative evaluation results in Table I and the difference

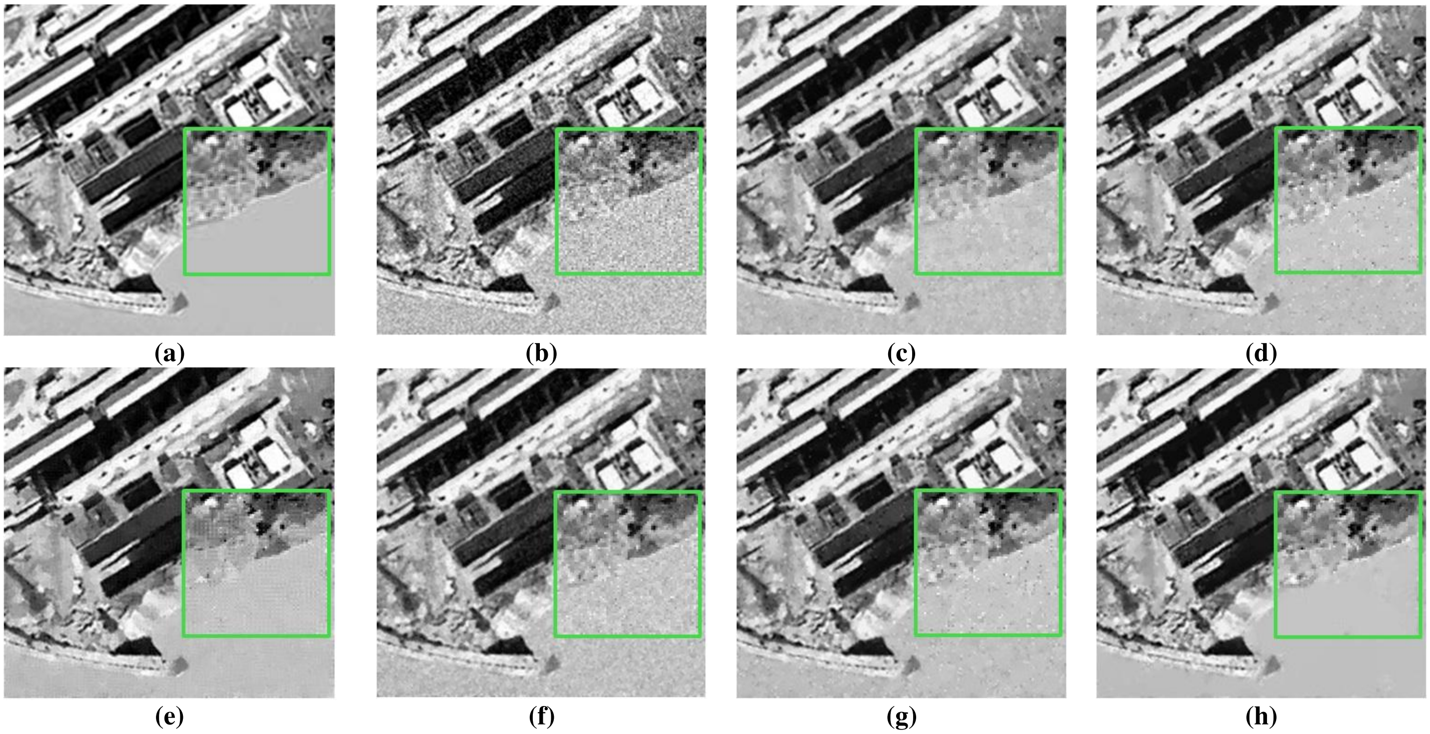


Fig. 8. A comparison of the denoising results using the different methods: (a) original image, (b) noisy image (noise variance = 18), (c) TV [26], (d) SATV [33], (e) SCAD [37], (f) ATV [34], (g) PSATV, and (h) RSATV.

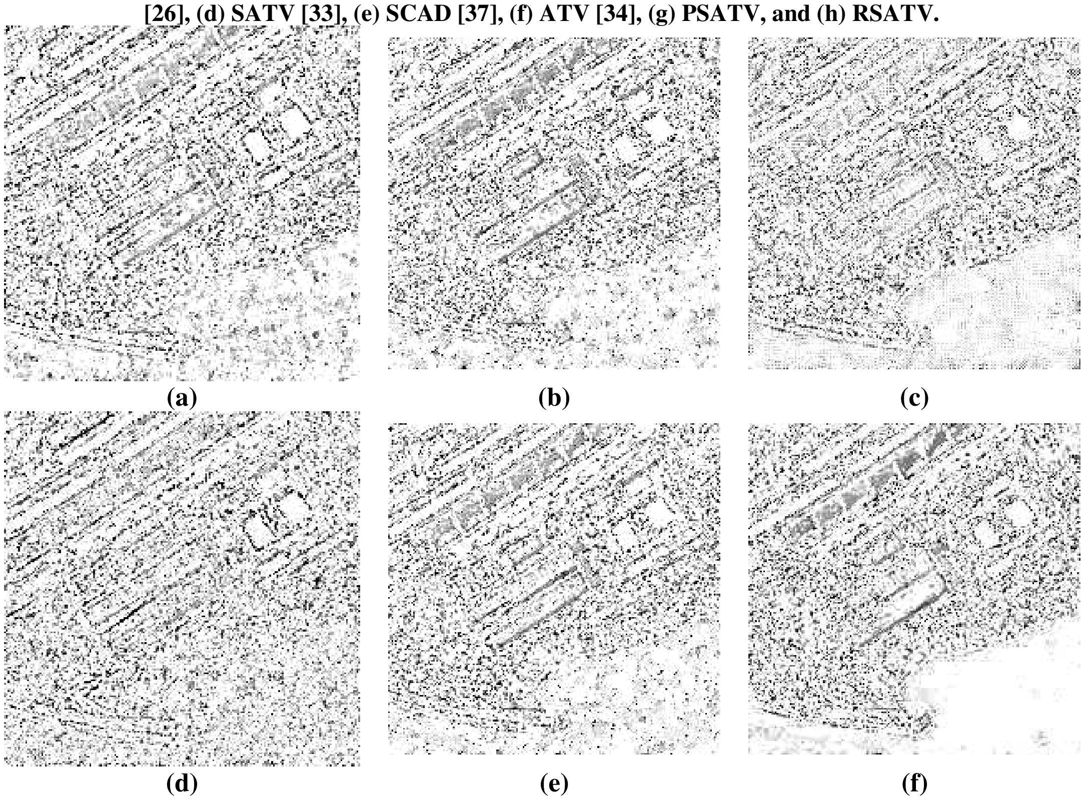


Fig. 9. The difference images between the denoising results and the ground truth image(intensities 2% linearly stretched): (a) TV, (b) SATV, (c) SCAD, (d) ATV, (e) PSATV, and (f) RSATV.

image between the denoising results and the ground truth image in Fig. 9. It can be seen that the RSATV model gives the highest PSNR and SSIM values at all the different noise intensities, which illustrates the noise robustness of the proposed model. Meanwhile, the difference image in Fig. 9 also illustrates that the proposed RSATV gives a better denoising result than the other five spatially adaptive TV models, especially in the flat regions.

TABLE I A QUANTITATIVE EVALUATION OF THE DENOISING RESULTS IN DIFFERENT NOISE CONDITIONS

Noise variance		TV [26]	SATV [33]	SCAD [37]	ATV [34]	PSATV	**RSATV**
8	PSNR	31.73	31.78	31.33	31.02	31.67	**32.14**
	SSIM	0.936	0.936	0.947	0.901	0.940	**0.964**
14	PSNR	27.94	27.77	26.97	27.20	27.64	**28.14**
	SSIM	0.887	0.893	0.879	0.834	0.870	**0.918**
18	PSNR	26.29	25.90	25.24	25.64	25.79	**26.46**
	SSIM	0.844	0.835	0.829	0.801	0.828	**0.897**

②Simulated Super-Resolution Experiment

Next, to assess the relative merits of the proposed methodology, we test it on an multi-frame image super-resolution problem. Firstly, it is tested on a simulated process.

In the simulated process, with the degradation model described in (1), the HR image is first shifted with sub-pixel displacements of (0,0), (0.5,0.5), (0.5,0) and (0,0.5) to produce four images. The image sequence is then convolved with a Gaussian-type PSF of 5×5 window size and unit variance, and downsampled with a factor of 2 in both the vertical and horizontal directions. Finally, zero-mean Gaussian noise added. We compare the proposed RSATV algorithm with the TV regularization in [26], the ATV model in [34], the PSATV model in in equation (11), and some other non-TV models used in the SR problem, including the Laplacian model in [13], and the BTV model in [21]. For the BTV model [21], the experiment parameters are: $P=2, \beta=20$, and $\alpha=0.8$. In all the prior models, the regularization parameter is adjusted until the best SR result is archived.

The SR results of the four simulated experimental datasets are shown in Fig. 10, which presents the SR results of the four HR images under noise variance 18. The difference images between the SR image and the true HR image are shown in Fig. 11. The quantitative evaluation results using the PSNR and SSIM indexes are shown in Table II-V.

From the SR results presented in the four figures, it can be seen that the proposed RSATV produces a better SR image than the TV model and the PSATV model. In the TV SR image, the noise in the flat regions of the image is not well suppressed, and some “pseudo-edges” are produced. When the noise intensity becomes higher, the “pseudo-edges” are more obvious. For the PSATV model, it was found that when the noise intensity is low, it can produce a better SR result than the TV model. However, when the noise intensity becomes higher, the SR result becomes worse, and, in the same way as with the TV model, some “pseudo-edges” are produced in the flat regions. The reason for this is that the PSATV model constructs the spatial information constraint from a pixel level, which causes the noise pixels in the flat regions to be falsely identified as edge pixels and given a small spatial weight. This causes the noise to be poorly suppressed and some “pseudo-edges” are produced. However, with the RSATV model, because the spatial information filtering and spatial weight clustering processes are added, a more accurate spatial constraint is enforced, and a better SR image is produced. The better performance can also be seen in the difference image, from which it can be seen that the RSATV model SR image is more close to the true HR image, especially in the flat regions.

The better performance of the proposed RSATV model can also be seen in the quantitative evaluation results presented in Tables II–V. It can be seen that the proposed approach produces the highest PSNR and SSIM values among the five models, which is consistent with the visual effect of the reconstructed images in Figs. 10 and 11.

③Real Data Super-Resolution Experiments

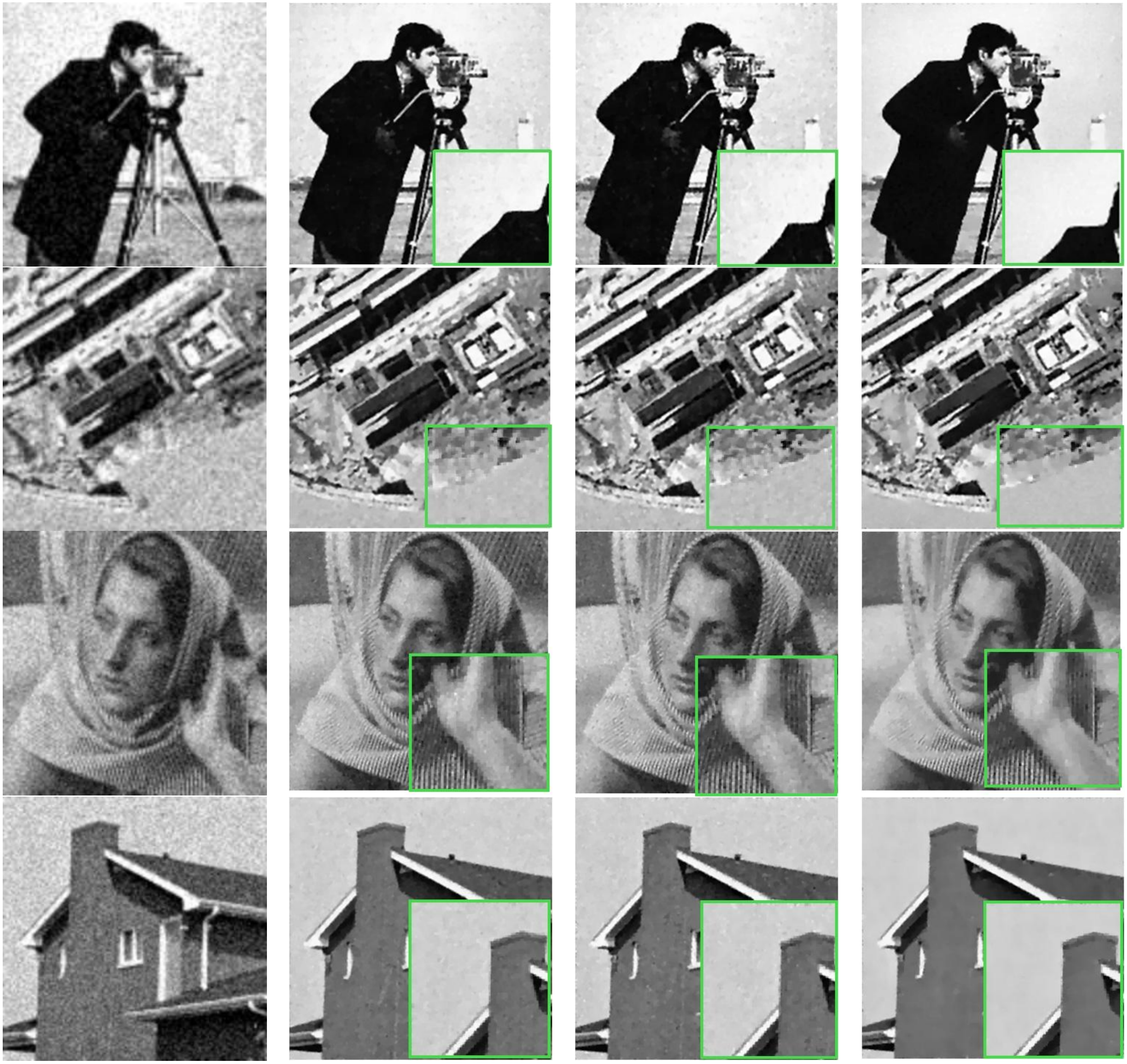
To verify the performance of the proposed RSATV model on real data, in Figs. 12–13, we present the SR results from the two real experiment datasets.

The experimental results for the “EIA” image sequence are presented in Figs. 12 and 13. The PSF of the sequence is assumed to be Gaussian-type, with a window size of 4×4 and a variance of 1. Of the five prior models, in the Laplacian model, the noise in the flat regions is not well suppressed, and in the BTV, TV and PSATV models, although the noise in the flat regions is suppressed, to some extent, some “artifacts” are produced. However, because the spatial information constraint is considered from a regional level, our RSATV model gives the most promising SR image. In the flat areas, the noise is well suppressed, but without losing the edge information, which can be clearly seen from the cropped regions presented in Fig. 13.

The results of the second real data experiment are shown in Figs. 14 and 15. The PSF of the sequence is assumed to be Gaussian-type, with a window size of 5×5 and a variance of 1. From the SR images, it can be clearly seen that the proposed RSATV model gives better SR results than the other five models. Because the spatial information constraint is considered from a regional level, and the spatial information filtering and spatial weight clustering processes can overcome the effect of the noise, the noise in the flat regions is well suppressed and the edge information is better preserved. However, in the TV model SR image, because the spatial information constraint is not considered, the noise in the flat regions is not well suppressed. In the PSATV SR result, because the spatial information constraint is considered just based on each pixel, it is more sensitive to noise in the flat regions, and a noise pixel is likely to be falsely identified as an edge pixel. This leads to the noise in the flat regions being poorly suppressed, which can be clearly seen in the cropped regions presented in Fig. 15.

C. Selection of the parameters in the RSATV model

Parameter β in (10): In (10), the parameter β is used to control the spatial information weighted parameter W_i in the SR process. If β is too small, the edge region pixels will be given almost the same spatial weight as the flat region pixels, and then the spatially weighted idea cannot be well realized. Conversely, if it is too large, a noise pixel in the flat regions will be given a small weight, and the noise cannot be effectively reduced. Therefore, in Tables VI–VIII, the effect of the parameter β on the final SR performance is analyzed. In all three Tables, it is found that a better SR image can be obtained when the parameter β is set at 0.05. Therefore, in this paper, parameter β is empirically set to be 0.05 in all the experiments. In our future research, we will explore some adaptive methods of selecting the optimal value of parameter β .



(a) (b) (c) (d)

Fig. 10. The super-resolution results of the "cameraman" image under a noise variance of 18: (a) the bilinear interpolation result, (b) the TV super-resolution image, (c) the PSATV super-resolution image, and (d) the proposed RSATV super-resolution image.

Parameter τ in (14): In (14), the spatial weight of the flat region μ_j is multiplied with a parameter τ to ensure that a large regularization strength is enforced and the noise is well suppressed. The setting of this parameter will affect the final SR performance. If it is too large, the SR result will be blurred, and, conversely, if it is too small, the noise cannot be suppressed well. Therefore, in this part, to analyze the effect of the parameter τ on the SR performance, using the "cameraman" image as an example, we plot the changes in the PSNR values using different τ values (from 1–10) under different noise intensities, which is shown in Fig. 16. From the plot, it can be seen that when the noise intensity is low, a small τ value should be used, and with the increase in the noise intensity, a large τ value is more suitable. From our test on the four simulated datasets, it is more appropriate to set the parameter τ to $1000\sigma^2$, where σ^2 is the additive noise variance, which is normalized to 0–1. For example, for the noise variances of 8, 11, 14, 16 and 18, which are respectively about 0.001, 0.002, 0.003, 0.004 and 0.005 when normalized to 0–1, the appropriate τ values are 1, 2, 3, 4 and 5, respectively. This is just an empirical example of the selection of the parameter τ, and from our experiments, it was found that a manual adjustment of the parameter is also possible and does not consume much time. We advise that, in most cases, the optimal value of this parameter will be

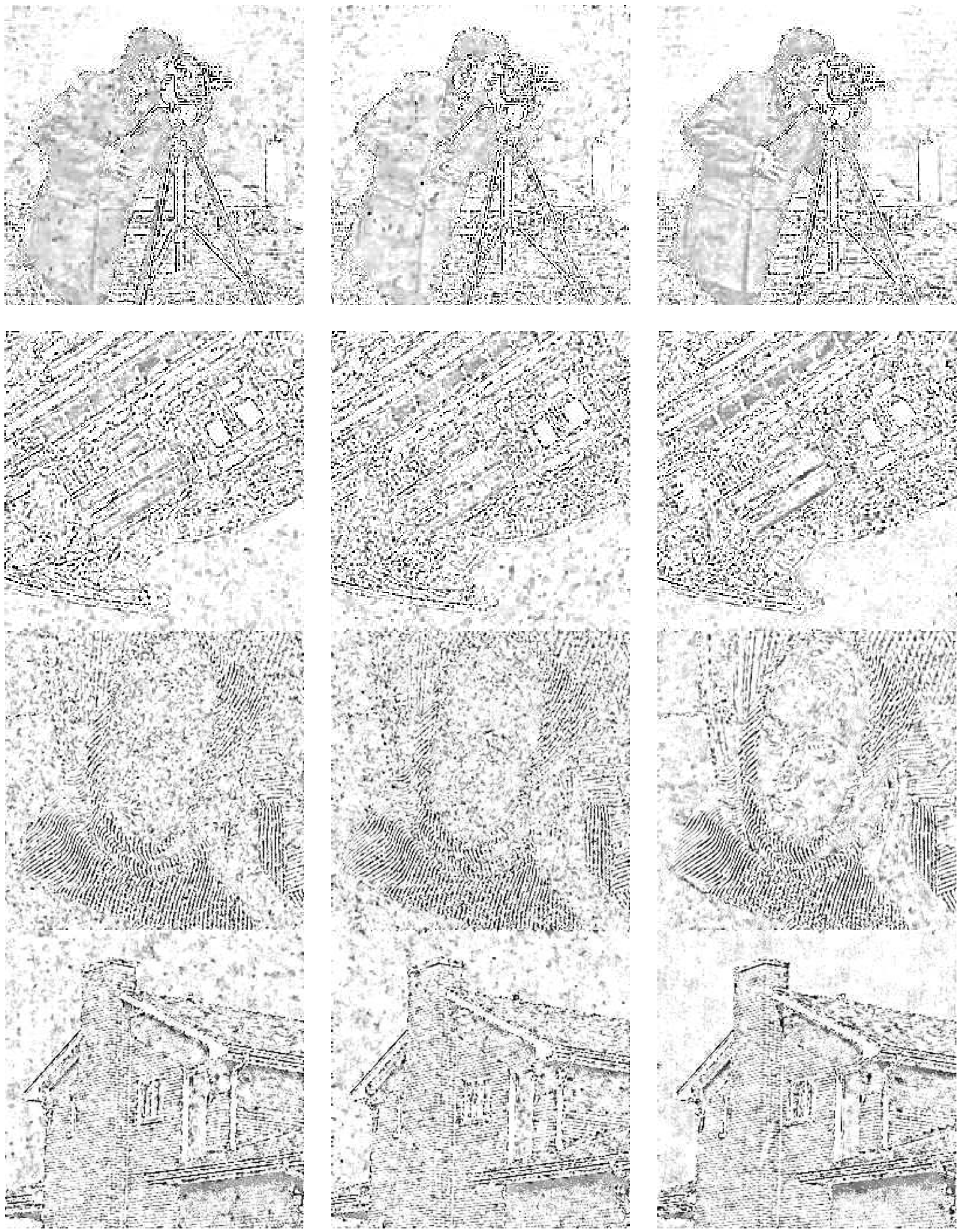


Fig. 11. The difference images between the SR results in Fig. 10 and the true HR image in Fig.7 (intensities 2% linearly stretched): (a) the TV super-resolution image (b) the PSATV super-resolution image, and (c) the proposed RSATV super-resolution image.

between 1 to 10. We will pay more attention to the adaptive setting of this parameter in our future research.

The cluster number n **in Section III-C:** In Section III-C, the k-means clustering method is adopted to construct the spatial weight from a regional perspective. In the k-means process, the cluster number n is an important parameter. Therefore, in this part, we present an analysis about the effect of the cluster number n on the final SR result. In Fig. 17, the change of the PSNR values with different cluster numbers of n (from 3–10) under different noise In conditions is plotted. In Fig.18, the change of the clustering results with different cluster numbers with the noise variance of 18 is also presented.

From the analysis, it can be seen that in low noise conditions, the change of the cluster number from 3 to 10 produces little effect on the SR result, while with an increasing noise intensity, the PSNR value changes a little, but it is not so obvious. In our experiments, the cluster number is empirically set to be 5 in all cases. In addition, it can be seen from Fig.18 that the structure information of the image should also be considered in the selection. For a simple structure image (such as the "house" image), the cluster number should be low, and for a complex structure image (for example, the "Barbara" image), the cluster number should be a little higher. In our future research, we will explore some adaptive methods of selecting the cluster number.

TABLE II. THE QUANTITATIVE EVALUATION RESULTS USING THE PSNR AND SSIM INDEXES OF THE "cameraman" IMAGE EXPERIMENT

Noise variance	Evaluation index	Laplacian[13]	BTV [21]	TV[26]	ATV [34]	PSATV	**RSATV**
8	PSNR	26.61	27.89	28.378	28.35	28.59	**28.87**
	SSIM	0.717	0.807	0.903	0.842	0.90753	**0.911**
11	PSNR	25.36	26.38	27.08	26.83	27.45	**27.78**
	SSIM	0.677	0.821	0.876	0.789	0.872	**0.887**
14	PSNR	24.56	25.55	26.38	26.04	26.67	**26.93**
	SSIM	0.674	0.786	0.855	0.748	0.852	**0.868**
16	PSNR	24.11	24.97	25.67	25.46	26.01	**26.28**
	SSIM	0.635	0.765	0.837	0.740	0.838	**0.853**
18	PSNR	23.70	24.55	25.56	24.98	25.62	**25.85**
	SSIM	0.635	0.743	0.804	0.735	0.808	**0.845**

TABLE III. THE QUANTITATIVE EVALUATION RESULTS USING THE PSNR AND SSIM INDEXES OF THE "aerial " IMAGE EXPERIMENT

Noise variance	Evaluation index	Laplacian[13]	BTV [21]	TV[26]	ATV [34]	PSATV	**RSATV**
8	PSNR	28.49	26.76	28.65	28.68	28.79	**29.02**
	SSIM	0.874	0.866	0.932	0.910	0.933	**0.936**
11	PSNR	26.70	25.61	27.16	27.01	27.26	**27.49**
	SSIM	0.817	0.860	0.901	0.863	0.903	**0.914**
14	PSNR	25.86	24.71	26.24	26.17	26.42	**26.55**
	SSIM	0.808	0.834	0.881	0.847	0.884	**0.903**
16	PSNR	25.21	24.07	25.37	25.47	25.63	**25.81**
	SSIM	0.781	0.808	0.863	0.827	0.858	**0.891**
18	PSNR	24.61	23.61	24.85	25.04	25.19	**25.22**
	SSIM	0.772	0.784	0.847	0.814	0.844	**0.857**

TABLE IV. THE QUANTITATIVE EVALUATION RESULTS USING THE PSNR AND SSIM INDEXES OF THE "Barbara" IMAGE EXPERIMENT

Noise variance	Evaluation index	Laplacian[13]	BTV [21]	TV[26]	ATV [34]	PSATV	**RSATV**
8	PSNR	29.97	28.17	30.32	30.17	30.35	**30.59**
	SSIM	0.826	0.804	0.880	0.868	0.890	**0.896**
11	PSNR	28.22	26.82	28.65	28.27	28.66	**28.83**
	SSIM	0.777	0.773	0.834	0.799	0.842	**0.855**
14	PSNR	27.13	25.93	27.65	27.23	27.68	**27.93**
	SSIM	0.754	0.728	0.805	0.783	0.811	**0.824**
16	PSNR	26.47	25.40	26.96	26.53	26.99	**27.19**
	SSIM	0.716	0.703	0.783	0.747	0.791	**0.799**
18	PSNR	25.89	24.87	26.54	26.04	26.50	**26.73**
	SSIM	0.701	0.672	0.753	0.726	0.757	**0.779**

TABLE V. THE QUANTITATIVE EVALUATION RESULTS USING THE PSNR AND SSIM INDEXES OF THE "house" IMAGE EXPERIMENT

Noise variance	Evaluation index	Laplacian[13]	BTV [21]	TV[26]	ATV [34]	PSA TV	**RSATV**
8	PSNR	30.78	30.73	32.01	31.65	32.08	**32.18**
	SSIM	0.800	0.809	0.837	0.821	0.839	**0.839**
11	PSNR	29.41	29.68	30.88	30.32	30.90	**31.06**
	SSIM	0.725	0.786	0.811	0.781	0.814	**0.817**
14	PSNR	28.63	28.99	30.10	29.42	30.18	**30.21**
	SSIM	0.714	0.766	0.794	0.748	0.797	**0.799**
16	PSNR	28.00	28.43	29.63	28.81	29.55	**29.68**
	SSIM	0.712	0.750	0.779	0.720	0.783	**0.783**
18	PSNR	27.60	28.05	29.19	28.37	29.08	**29.23**
	SSIM	0.689	0.738	0.775	0.706	0.770	**0.779**

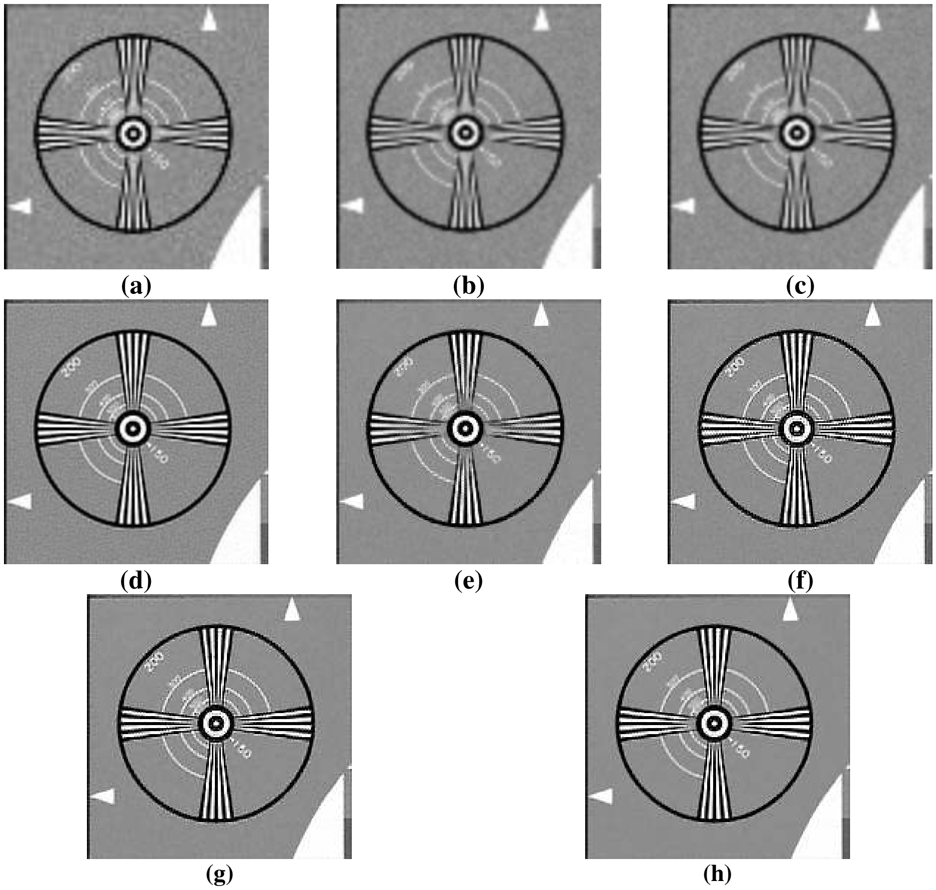


Fig. 12. Reconstruction results of the "EIA" image sequence: (a) LR image, (b) bilinear interpolation, (c) bicubic interpolation, (d) Laplacian regularization, (e) BTV regularization, (f) TV regularization, (g) PSATV regularization, and (h) RSATV regularization.

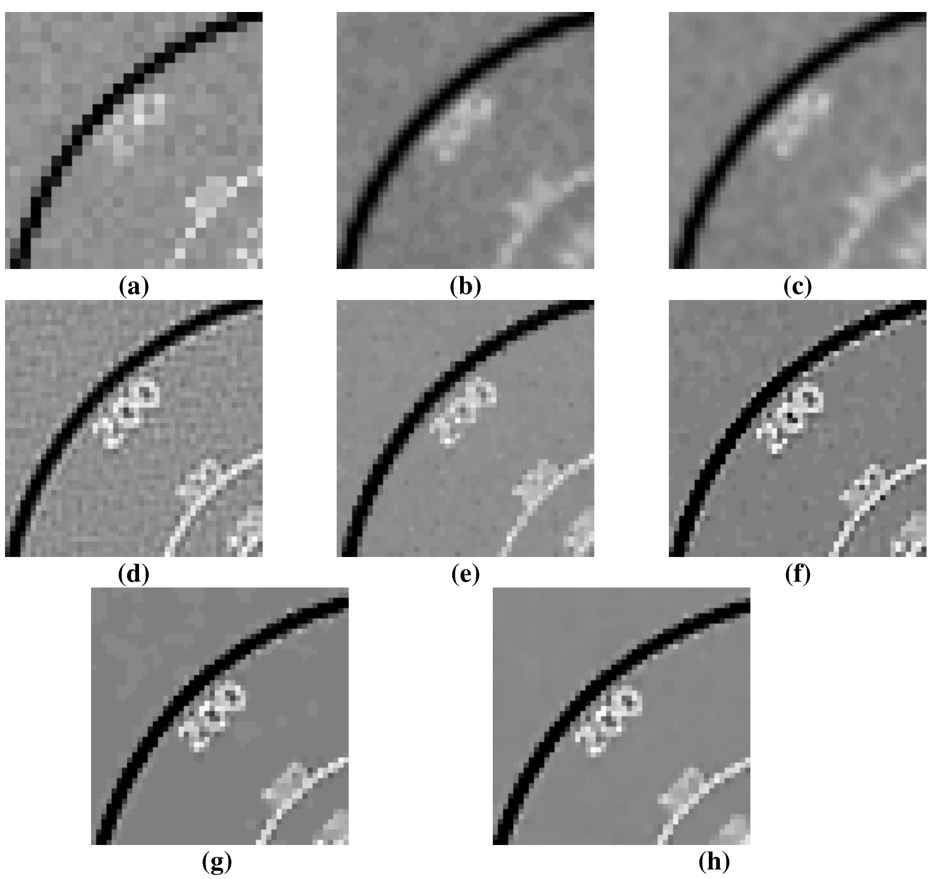


Fig. 13. Detail regions cropped from Fig. 12: (a) LR image, (b) bilinear interpolation, (c) bicubic interpolation, (d) Laplacian regularization, (e) BTV regularization, (f) TV regularization, (g) PSATV regularization, and (h) RSATV regularization.

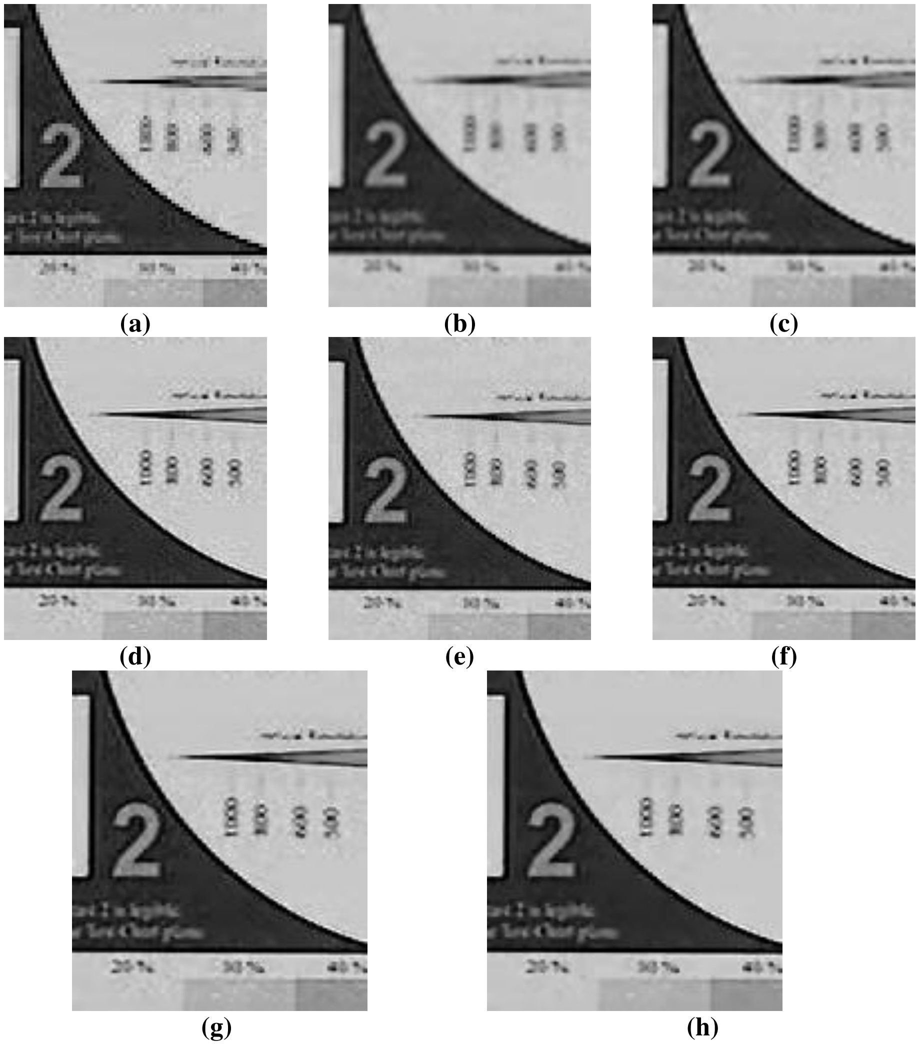
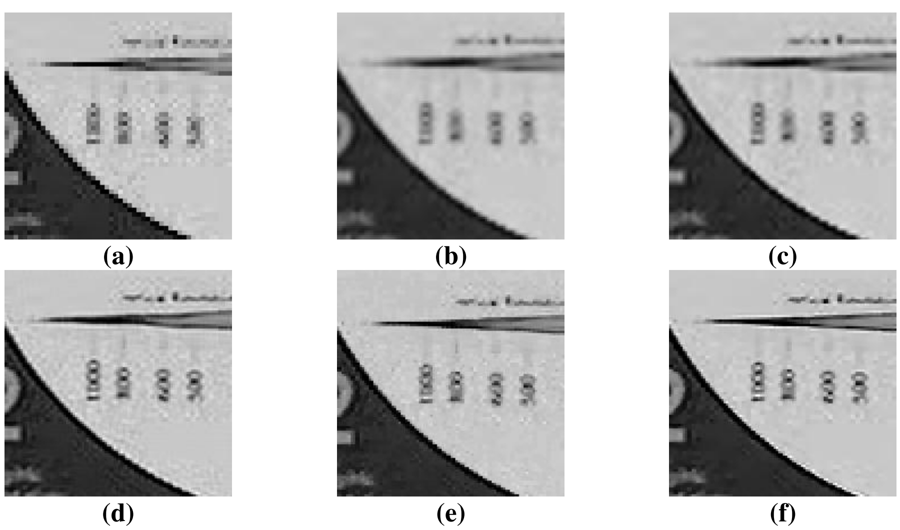


Fig. 14. Reconstruction results of the "surveillance" video sequence: (a) LR image, (b) bilinear interpolation, (c) bicubic interpolation, (d) Laplacian regularization, (e) BTV regularization, (f) TV regularization, (g) PSATV regularization, and (h) RSATV regularization.



(g) (h)

Fig 15. Detail regions cropped from Fig. 14: (a) LR image, (b) bilinear interpolation, (c) bicubic interpolation, (d) Laplacian regularization, (e) BTV regularization, (f) TV regularization, (g) PSATV regularization, and (h) RSATV regularization.

TABLE VI. THE CHANGE OF PSNR VALUES WITH DIFFERENT VALUES OF β UNDER DIFFERENT NOISE VARIANCES WITH THE "CAMERAMAN" IMAGE

β / Noise	0.001	0.005	0.01	**0.05**	0.1	0.5	1
8	28.13	28.51	28.51	**28.87**	28.77	28.27	28.20
14	26.10	26.67	26.74	**26.93**	26.11	25.65	25.51
18	25.19	25.37	25.23	**25.85**	25.37	24.22	24.12

TABLE VII. THE CHANGE OF PSNR VALUES WITH DIFFERENT VALUES OF β UNDER DIFFERENT NOISE VARIANCES WITH THE "AERIAL" IMAGE

Noise	0.001	0.005	0.01	**0.05**	0.1	0.5	1
8	28.67	28.61	28.88	**29.02**	28.55	28.40	28.49
14	26.14	26.46	26.26	**26.55**	25.67	25.04	24.99
18	23.52	24.99	25.04	**25.22**	24.85	24.55	24.30

TABLE VIII. THE CHANGE OF PSNR VALUES WITH DIFFERENT VALUES OF β UNDER DIFFERENT NOISE VARIANCES WITH THE "BARBARA" IMAGE

β / Noise	0.001	0.005	0.01	**0.05**	0.1	0.5	1
8	29.95	30.14	30.41	**30.59**	30.41	30.08	29.85
14	26.97	27.14	27.46	**27.93**	27.65	26.44	26.25
18	25.56	25.29	26.19	**26.73**	26.63	25.35	25.15

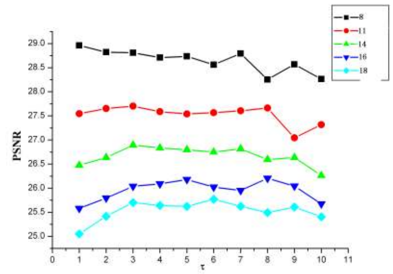


Fig. 16. Change in the SR performance with different τ values under different noise variances.

VI CONCLUSION

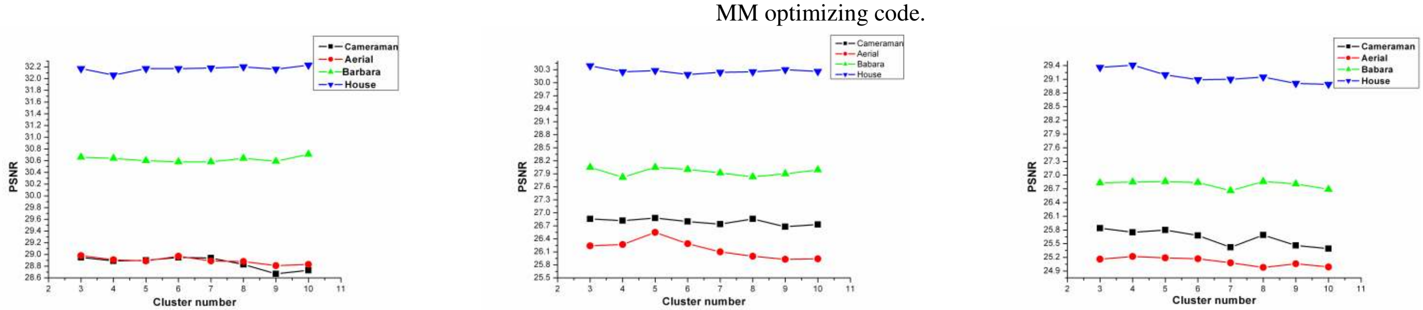
The traditional spatially adaptive total variation model has the shortcoming of being sensitive to noise, and it performs poorly in high noise intensity conditions. To overcome this, in this paper, we propose a regional spatially adaptive total variation (RSATV) super-resolution algorithm with spatial information filtering and clustering. The spatial

information is first extracted for each pixel, and then the spatial information filtering process and spatial weight clustering process are added. With these two processes, the regularization strength of the total variation model is adjusted for each region with different spatial properties, rather than for each pixel, as in the traditional spatially adaptive TV model. The simulated and real data experiments presented in Section V show that the proposed RSATV model can better suppress the noise in the flat regions of an image, without losing the edge and detail information.

In our future research, we will focus on adaptive parameter selection for the method, and we will also investigate the use of more efficient optimization algorithms to accelerate the solution speed of the RSATV model, such as the FISTA and MFISTA algorithms detailed in [53]–[55], Furthermore, some noise-robust spatial feature indicators, such as steering weights [56], will also be considered, to further improve the spatial weight construction process of the proposed algorithm.

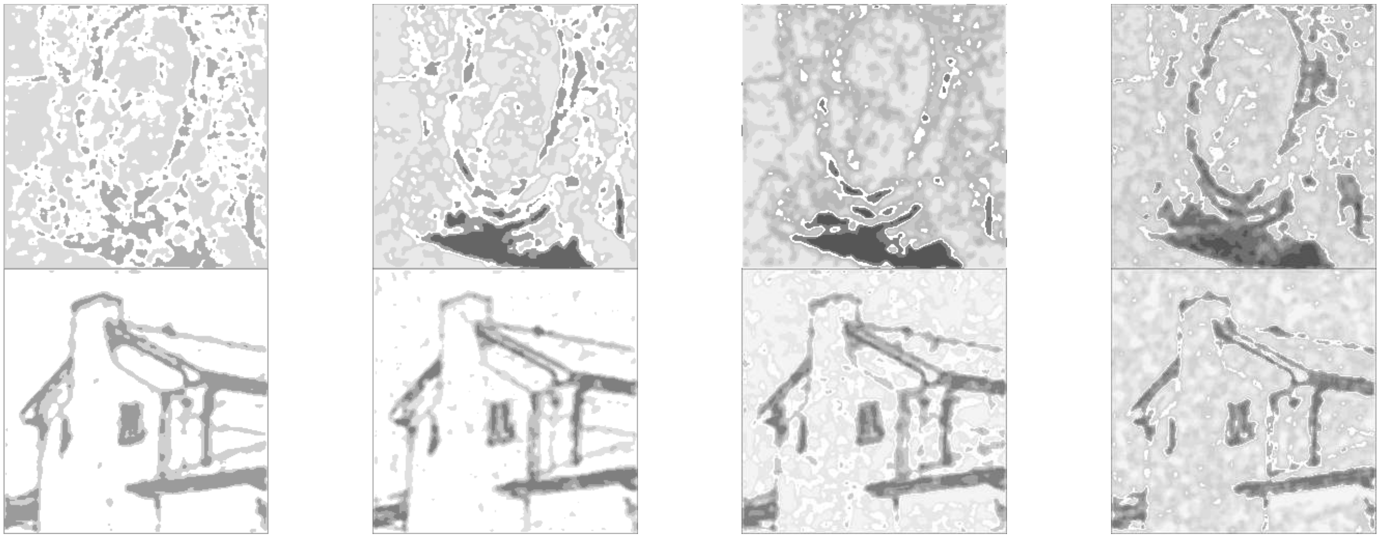
ACKNOWLEDGMENTS

The authors would like to thank the work of the editor and the anonymous reviewers. We also would like to thank Professor Peyman Milanfar and Michael Elad for providing the BTV SR code and the real experiment data, and Professor J. Bioucas-Dias for providing the MM optimizing code.



(a) (b) (c)

Fig. 17. The effect of the cluster number n on the SR performance: (a) noise variance = 8, (b) noise variance = 14, and (c) noise variance = 18.



(a) (b) (c) (d)

Fig. 18. Change of the clustering results with different clustering numbers for the spatial weight (a) cluster number=3,(b) cluster number=5,(c) cluster number=7, and (d) cluster number =10;

REFERENCES

[1] H. Greenspan, "Super-Resolution in Medical Imaging," Comput. J., vol. 52, no. 1, pp. 43–63, 2009.

[2] X. Huang, L. Zhang, and P. Li, "Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery," IEEE Trans. Geosci. Remote Sens. Lett., vol. 4, no. 2, pp. 260–264, 2007.

[3] X. Huang and L. Zhang, "An Adaptive Mean-Shift Analysis Approach for Object Extraction and Classification from Urban Hyperspectral Imagery," IEEE Trans. Geosci. Remote Sens., vol. 46, no. 12, pp. 4173–4185, Dec. 2008.

[4] L. Zhang, H. Zhang, H. Shen, and P. Li, "A super-resolution reconstruction algorithm for surveillance images," Signal Process., vol. 90, no. 3, pp. 848–859, 2010.

[5] R. Tsai and T. Huang, "Multiple frame image restoration and registration," in Advances in Computer Vision and Image Processing, vol. 1, no. 2, pp. 317–339, 1984.

[6] S. Kim, N. Bose, and H. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," IEEE Trans. Acoust., Speech, Signal Process., vol. 38, no. 6, pp. 1013–1027, 1990.

[7] S. Kim and W. Su, "Recursive high-resolution reconstruction of blurred multiframe images," IEEE Trans. Image Process., vol. 2, no. 4, pp. 534–539, 1993.

[8] H. Ur and D. Gross, "Improved resolution from sub-pixel shifted pictures," Comput. Vision Graph.: Graphical Models and Image Processing, vol. 54, no. 2, pp. 181–186, 1992.

[9] M. Alam, J. Bognar, R. Hardie, and B. Yasuda, "Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames," IEEE Trans. Instrum. Meas., vol. 49, no. 5, pp. 915–923, 2000.

[10]B. Tom and A. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images," in Proc. 1995 IEEE Int. Conf. Image Processing, vol. 2, Washington, DC, pp. 539–542.

[11] R. Schultz and R. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process., vol. 5, no. 6, pp. 996–1011, 1996.

[12] S. Belekos, N. Galatsanos, and A. Katsaggelos, "Maximum a Posteriori Video Super-Resolution Using a New Multichannel Image Prior," IEEE Trans. Image Process., vol. 19, no. 6, pp. 1451–1464, 2010.

[13] R. Hardie, K. Barnard, and E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1621–1633, 1997.

[14] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP Approach for Joint Motion Estimation, Segmentation and Super-resolution," IEEE Trans. Image Process., vol. 16, no. 2, pp. 479–490, 2007.

[15] M. Elad and A. Feuer, "Restoration of a single super resolution image from several blurred, noisy, and undersampled measured images," IEEE Trans. Image Process., vol. 6, no. 12, pp. 1646–1658, 1997.

[16] M. Protter, M. Elad, H. Takeda, and P. Milanfar "Generalizing the Nonlocal-Means to Super-Resolution Reconstruction," IEEE Trans. Image Process., vol. 18, no. 1, pp. 36–51, 2009.

[17] H. Takeda, P. Milanfar, M. Protter, and M. Elad, "Super-Resolution Without Explicit Subpixel Motion Estimation," IEEE Trans. Image Process., vol. 18, no. 9, pp. 1958–1975, 2009.

[18] N. Nguyen and P. Milanfar, "A wavelet-based interpolation restoration method for superresolution (wavelet superresolution)," Circuits, Systems, and Signal Processing, vol. 19, no. 4, pp. 321–338, 2000.

[19] H. Ji and C. Fermuller, "Robust Wavelet-Based Super-Resolution Reconstruction: Theory and Algorithm," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp. 649–660, 2009.

[20] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," IEEE Signal Processing Mag., vol. 20, no. 3, pp. 21–36, 2003.

[21] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super-resolution," IEEE Trans. Image Process., vol. 13, no. 10, pp. 1327–1344, 2004.

[22] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," Int. J. Imag. Syst. Tech., vol. 14, no. 2, pp. 47–57, 2004.

[23] A. K. Katsaggelos, R. Molina, and J. Mateos, Super Resolution of Images and Video. 1st ed., Morgan and Claypool, 2007

[24] P. Milanfar, Super Resolution imaging. 1st ed., CRC Press, 2010

[25] L. Rudin, S. Osher, and E Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D: Nonlinear Phenomena, vol. 60, no. 1, pp. 259–268, 1992.

[26] M. Ng, H. Shen, E. Lam, and L. Zhang. "A total variation regularization based super-resolution reconstruction algorithm for digital video," EURASIP J. Adv. Sig. Pr., Article ID 74585, pp. 1–16, 2007.

[27] A. Marquina and S. Osher, "Image Super-Resolution by TV-Regularization and Bregman Iteration," J. Sci. Comput., vol. 37, no. 3, pp. 367–382, 2008.

[28] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image Super-Resolution Via Sparse Representation," IEEE Trans. Image Process., vol.19, no.11, pp. 2861–2873, Nov. 2010.

[29] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and supper-resolution by adaptive sparse domain selection and adaptive regularization," IEEE Trans. Image Process., vol. 20, no. 7, pp. 1838–1857, Jul. 2011.

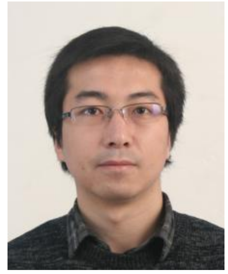
[30] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based Image Denoising via Dictionary Learning and Structural Clustering," in Proc. 2011 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2011), 20–25 Jun. 2011, pp. 457–464.

[31] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image deblurring with locally adaptive and non locally robust regularization," in Proc. 2011 IEEE Int. Conf. Image Processing (ICIP2011).

[32] T. Chan, S. Esedoglu, F. Park, A. Yip, Recent developments in total variation image restoration, in: N. Paragios, Y. Chen, O. Faugeras (Eds.), Handbook of Mathematical Models in Computer Vision, Springer, Berlin, 2004.

[33] D. M. Strong, P. Blomgren, and T. F. Chan, "Spatially adaptive local feature-driven total variation minimizing image restoration," in: F. Preteux, J. L. Davidson, E. R. Dougherty (Eds.), Conf. Statistical and Stochastic Methods in Image Processing II, San Diego, CA, 1997, pp. 222–233.

[34] Q. Chen, P. Montesinos, Q. Sun, P. Heng, and D. Xia, "Adaptive total variation denoising based on difference curvature, " Image Vision Comput., vol. 28, no. 3, pp. 298–306, 2010.
[35] W. Guo and F. Huang, "Adaptive total variation based filtering for MRI images with spatially inhomogeneous noise and artifacts", ISBI 2009,pp: 101-104
[36] G. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational Bayesian Image Restoration With a Product of Spatially Weighted Total Variation Image Priors," IEEE Trans. Image Process., vol. 19, no. 2, pp. 351–362, Feb. 2010.
[37] A. Chopra and H. Lian, "Total variation, adaptive total variation and nonconvex smoothly clipped absolute deviation penalty for denoising blocky images," Pattern Recogn., vol. 43, no. 8, pp. 2609–2619, 2010.
[38] T. Saito and T. Komatsu, "Super-resolution Sharpening-Demosaicking with Spatially Adaptive Total-Variation Image Regularization", Lecture Notes in Computer Science, 2005, Volume 3767/2005, 246-256 2005.
[39] Q. Yuan, L. Zhang, H. Shen, and P. Li, "Adaptive multiple-frame image super-resolution based on U-curve," IEEE Trans. Image Process., vol. 19, no. 12, pp. 3157–3170, 2010.
[40] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," Inverse Probl., vol. 19, pp. 165–187, 2003.
[41]J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," SIAM J. Imag. Sci., vol. 2, no. 2, pp. 569–592, May 2009.
[42] P. Chatterjee and P. Milanfar, "Clustering-based Denoising with Locally Learned Dictionaries," IEEE Trans. Image Process., vol. 18, no. 7, pp. 1438–1451, Jul. 2009.
[43] P. Chatterjee and P. Milanfar, "Patch-based Near-optimal Image Denoising," IEEE Trans. Image Process., vol. 21, no. 4, pp. 1635–1649, Apr. 2012.
[44]J. Bioucas-Dias, M. Figueiredo, and J. Oliveira, "Total variation image deconvolution: A majorization-minimization approach," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing – ICASSP 2006, Toulouse, France, May 2006.
[45] D. Hunter and K. Lange, "A tutorial on MM algorithms," Am. Stat., vol. 58, no. 2, pp. 30–37, Feb. 2004.
[46]J. Bioucas-Dias, M. Figueiredo, and J .Oliveira, "Adaptive total-variation image deconvolution: A majorization-minimization approach," in Proc. EUSIPCO 2006, Florence, Italy, Sept. 2006.
[47] J. Oliveira, J. Bioucas-Dias, and M. Figueiredo, "Adaptive total variation image deblurring: A majorization-minimization approach," Signal Process., vol. 89, no. 9, pp. 1683–1693, 2009.
[48] P. Rodriguez and B. Wohlberg, "Efficient Minimization Method for a Generalized Total Variation Functional," IEEE Trans. Image Process., vol. 18, no. 2, pp. 322–332, Feb. 2009.
[49] http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html
[50] P. Vandewalle, S. Susstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," EURASIP J. Adv. Sig. Pr., no. 71459, pp. 1–14, 2006.
[51] Z. Wang, A. Bovik, and H. Sheikh, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004.
[52] Z. Wang and A. Bovik, "Mean squared error: love it or leave it? - A new look at signal fidelity measures," IEEE Signal Processing Mag., vol. 26, no. 1, pp. 98–117, Jan. 2009.
[53] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imag. Sci., vol. 2, no. 1, pp. 183–202, 2009.
[54] A. Beck and M. Teboulle, "Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems," IEEE Trans. Image Process., vol. 18, no.11, pp. 2419–2434, Nov. 2009.
[55] Y. E. Nesterov, Gradient Methods for Minimizing Composite Objective Function Tech. report, CORE, Sept. 2007 [Online]. Available: http://www.ecore.beDPs/dpll91313936.pdf
[56] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," IEEE Trans. Image Process., vol. 16, no. 2, pp. 349–366, Feb. 2007.



Qiangqiang Yuan received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively. In July 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently an Assistant Professor. His research interests focus on image reconstruction, remote sensing image processing and application, data fusion and assimilation.

Dr. Yuan is a Reviewer of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING, and IEEE SINGNAL PROCESSING LETTER.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education of China. He is currently a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 260 research papers. He is the holder of five patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a Fellow of the IEE, Executive Member (Board of Governor) of the China National Committee of International Geosphere–Biosphere Programme, Executive Member of the China Society of Image and Graphics, etc. He regularly serves as a General Chair of 4th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing(Whispers 2012), Cochair of the series SPIE Conferences on Multispectral Image Processing and Pattern Recognition, Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and Geoinformatics Symposiums. He also serves as an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, International Journal of Ambient Computing and Intelligence, International Journal of Image and Graphics, International Journal of Digital Multimedia Broadcasting, Journal of Geo-spatial Information Science, and Journal of Remote Sensing.

Huanfeng Shen (M'11) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University,Wuhan, China, in 2002 and 2007,respectively.

In July 2007, he joined the School of Resources and Environmental Science, Wuhan University, where he is currently a Full Professor. His research interests include image processing (for quality improvement), remote sensing application, data fusion, and assimilation. He has published more than 60 research papers. He has been supported by several talent programs, including the New Century Excellent Talents by the Ministry of Education of China (2011) and the Hubei Science Fund for Distinguished Young Scholars (2011).