

Iterative Support Detection-Based Split Bregman Method for Wavelet Frame-Based Image Inpainting

Liangtian He, *Member, IEEE*, and Yilun Wang, *Member, IEEE*

Abstract—The wavelet frame systems have been extensively studied due to their capability of sparsely approximating piecewise smooth functions, such as images, and the corresponding wavelet frame-based image restoration models are mostly based on the penalization of the ℓ_1 norm of wavelet frame coefficients for sparsity enforcement. In this paper, we focus on the image inpainting problem based on the wavelet frame, propose a weighted sparse restoration model, and develop a corresponding efficient algorithm. The new algorithm combines the idea of iterative support detection method, first proposed by Wang and Yin for sparse signal reconstruction, and the split Bregman method for wavelet frame ℓ_1 model of image inpainting, and more important, naturally makes use of the specific multilevel structure of the wavelet frame coefficients to enhance the recovery quality. This new algorithm can be considered as the incorporation of prior structural information of the wavelet frame coefficients into the traditional ℓ_1 model. Our numerical experiments show that the proposed method is superior to the original split Bregman method for wavelet frame-based ℓ_1 norm image inpainting model as well as some typical ℓ_p ($0 \leq p < 1$) norm-based nonconvex algorithms such as mean doubly augmented Lagrangian method, in terms of better preservation of sharp edges, due to their failing to make use of the structure of the wavelet frame coefficients.

Index Terms—Image inpainting, iterative support detection, augmented Lagrangian, split Bregman, wavelet frames, sparse optimization, nonconvex optimization.

I. INTRODUCTION

INPAINTING refers to problems of filling-in the missing parts in images, and it arises, for examples, from removing scratches in photos, restoring ancient drawings, and filling in the missing pixels of images transmitted through noisy channels. For the simplicity of notations, we denote the image as vectors in \mathbb{R}^n with n being equal to the number of total pixels. Image inpainting is often formulated as a linear inverse problem:

$$f = Au + \eta, \quad (1)$$

where f is the observed image, η denotes the additive Gaussian white noise with mean being 0 and variance

Manuscript received November 5, 2013; revised May 2, 2014 and September 16, 2014; accepted September 18, 2014. Date of publication October 8, 2014; date of current version November 12, 2014. This work was supported in part by the National Science Foundation of China under Grant 11201054 and Grant 91330201 and in part by the Fundamental Research Funds for the Central Universities under Grant ZYGX2012J118 and Grant ZYGX2013Z005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ali Bilgin.

The authors are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: 970754953@qq.com; yilun.wang@gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2014.2362051

being σ^2 , the linear operator A is a projection operator, or more precisely, a diagonal matrix with diagonals 1 if the corresponding pixel is known, or 0 otherwise, and u is the unknown true image.

To restore u from f , one of the most natural choices is solving the following least square problem

$$\min_{u \in \mathbb{R}^n} \|Au - f\|_2^2 \quad (2)$$

where $\|\cdot\|_2$ denotes the ℓ_2 norm. However, the noise η possessed by f will be amplified greatly, since the matrix A is usually ill-conditioned. Therefore, to suppress the effect of noise and also preserve the sharp attributes of images, e.g., edges and boundaries, various regularization based methods have been proposed to obtain a reasonable approximated solution from (1). Among them, variational methods and wavelet frames based approaches are widely used and have been proven successful.

The trend of variational methods for image processing begins with the refined Rudin-Osher-Fatemi (ROF) model [11] which penalized the total variation (TV) of u , due to its advantages in exploiting blocky image structures and preserving image edges, and the ROF model is particularly effective on restoring images that are piecewise constant. In the past few years, many corresponding efficient algorithms have been proposed to further promote the applicability of the ROF model, including our proposed operator splitting algorithm [55], and many others [56], [57] and el al. In addition, many other types of variational models have been also proposed afterward and we refer the interested readers to [12], [23]–[27], and the references therein for further studies.

Wavelet frame based methods are relatively new and their basic idea is that images can be sparsely approximated by properly designed wavelet frames. Correspondingly, the commonly regularization method for these models is the ℓ_1 -norm of the wavelet frame coefficients due to its sparsity enforcement and convexity. Although the forms of wavelet frame are similar to variational methods, they were considered as different approaches, partially because the wavelet frame models are defined for discrete data, while the variational methods assume continuity. While some studies in the literatures [54] indicated that there was a relationship between haar wavelet and variational methods, it is still not clear whether there exist general relations between wavelet frames and variational models in the context of image restoration, though a recent paper [30] shows that, for one of the wavelet frame based models namely analysis based approach, it can be viewed as a finite difference approximation of a certain type of

general variational model, and the approximation will be exact if image resolution goes to infinity. In addition, it has been shown in [1], [2], [6], and [28]–[30] that the wavelet frame based models are often superior than some variational models such as ROF model because the multiresolution structure and redundancy property of wavelet frames allow to adaptively choose proper differential operators according to the order of the singularity of the underlying solutions for different regions of a given image [30].

The main contributions of this paper are summarized as follows. We propose a new wavelet frame based weighted ℓ_1 minimization model for image inpainting. While the new model is non-convex, we give an efficient multi-stage convex relaxation algorithm based on the ideas of iterative support detection [13] and alternative optimization to find a reasonable good solution, which is not necessarily the global solution though. More important, when performing the key component, i.e. support detection, we make use of the multi-level structure of wavelet frame coefficients to improve the support detection quality rather than the plain thresholding adopted in the original implementation proposed in [13], and the better performance is achieved than the plan ℓ_1 model and the $\ell_p(0 \leq p < 1)$ model which both fail to make use of the structure information of the wavelet coefficients. Numerical experiments empirically demonstrate the advantages of this new algorithm in terms of better preserving of image structures such as edges.

The rest of this paper is organized as follows. In the next section, we first introduce wavelet frames and a corresponding wavelet frame based ℓ_1 norm image inpainting model. An efficient algorithm proposed by [2], [30], and [41] is then reviewed. In Section III, we introduced our new wavelet frame based weighted ℓ_1 image inpainting model and proposed a corresponding effective algorithm motivated by the idea of the iterative support detection. In Section IV, we compare the performance of our new method with the ℓ_1 norm based approach proposed by [30] and $\ell_p(0 \leq p < 1)$ norm based approach [7]. Section V is devoted to the conclusion of this paper and discussions on some possible future research.

II. BACKGROUNDS

A. Wavelet Frame and Wavelet Frame Based Models

We first briefly introduce the concepts of tight frame and tight wavelet frame, then review some of the typical wavelet restoration models. Interested readers can consult [31], [32], and [5], [6], [33] for further detailed introduction of wavelet frame and its applications.

A countable set $X \subset L_2(\mathbb{R})$ is called a tight frame of $L_2(\mathbb{R})$ if

$$f = \sum_{h \in X} \langle f, h \rangle h, \quad \forall f \in L_2(\mathbb{R}).$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of $L_2(\mathbb{R})$. The tight frame X is called a tight wavelet frame if the elements of X are generated by dilations and translations of finitely many functions called framelets. The construction

of framelets can be obtained according to the unitary extension principle (UEP) [31]. Following the common experiment implementations, we will also use the piecewise linear B-spline framelets constructed by considering the balance of the time and quality. Given a 1D framelet system for $L_2(\mathbb{R})$, the s -dimensional tight wavelet frame system for $L_2(\mathbb{R}^s)$ can be easily constructed by using tensor products of 1D framelets [6], [32].

In the discrete setting, we will use $W \in \mathbb{R}^{m \times n}$ with $m \geq n$ to denote fast tensor product framelet decomposition and use W^T to denote the fast reconstruction. Then according to the unitary extension principle [31], we have $W^T W = I$, and then we will have a perfect reconstruction i.e., $u = W^T W u$ for any image u . We further denote an L -level framelet decomposition of u as

$$Wu = (\dots, W_{l,j}u, \dots) \quad \text{for } 0 \leq l \leq L-1, j \in \mathcal{I}$$

where \mathcal{I} denotes the index set of all framelet bands and $W_{l,j}u \in \mathbb{R}^n$ is the wavelet frame coefficients of u in bands j at level l . We will also use $\alpha \in \mathbb{R}^m$ to denote the frame coefficients, i.e., $\alpha = Wu$, where

$$\alpha = (\dots, \alpha_{l,j}, \dots), \quad \text{with } \alpha_{l,j} = W_{l,j}u.$$

Since wavelet frame systems are redundant systems, it is clear that the mapping from the image u to its coefficients is not one-to-one, i.e., the representation of u in the wavelet domain is not unique. There have existed several different wavelet frame based models proposed in the literature utilizing the sparseness of the wavelet frame coefficients, including the synthesis based approach [34]–[38], the analysis based approach [2]–[4], and the balanced approach [1], [39], [40].

In this paper, we will take the analysis based approach (shorten as ASBA) as the example to demonstrate our new idea, though it can also be applied to other wavelet frame based sparse models. We first review the analysis based approach for solving the image inpainting problem (1).

$$(ASBA) \quad \min_u \frac{1}{2} \|Au - f\|_2^2 + \|\lambda \cdot Wu\|_{1,q} \quad (3)$$

where $q = 1$ or $q = 2$ corresponding to anisotropic ℓ_1 norm and isotropic ℓ_1 norm, respectively. The ℓ_1 -norm can be extended to a generalized ℓ_1 -norm defined as

$$\|\lambda \cdot Wu\|_{1,q} = \left\| \sum_{l=0}^{L-1} \left(\sum_{j \in \mathcal{I}} \lambda_{l,j} |W_{l,j}u|^q \right)^{1/q} \right\|_1 \quad (4)$$

where $|\cdot|^p$ and $(\cdot)^{\frac{1}{p}}$ are entrywise operations. If we let $\alpha = Wu$ and substitute it into (3), we can get the rewritten form of (3) as follows

$$\min_{u,\alpha} \frac{1}{2} \|Au - f\|_2^2 + \|\lambda \cdot \alpha\|_{1,q} \quad s.t. \quad \alpha = Wu \quad (5)$$

it is obvious that (3) and (5) are equivalent.

In order to avoid confusion, it is necessary to explain the parameter λ and wavelet frame transform Wu in more details. λ and Wu are 3-mode tensors. The first mode of λ and Wu denotes the level, e.g., framelet decomposition level to be 4 (i.e., $L = 4$) in our numerical tests. The second mode of

λ and Wu denotes pixel location, i.e., the number of total pixels is n . The third mode of λ and Wu denotes the bands, e.g., the number of bands in given level at a given pixel is 9 if we use the piecewise linear B-spline framelets in this paper. We denote the elements of λ and Wu as $\lambda_{l,i,j}$, $W_{l,i,j}u$, respectively, where $0 \leq l \leq L-1$, $i \in \{1, 2, \dots, n\}$, $j \in \mathcal{I}$. Specifically, if the subscripts only contain two of (l, i, j) , the elements of the omitting mode constitute a vector e.g., $W_{l,j}u \in \mathbb{R}^n$ denotes a vector, where the elements of it are the wavelet frame coefficients of u in bands j at level l , $\lambda_{l,j} \in \mathbb{R}^n$ denotes a vector. The elements of it are the $\lambda_{l,i,j}$ in bands j at level l , $\lambda_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ denotes a vector similarly, where the elements of it are the $\lambda_{l,i,j}$ in level l at pixel i .

Note that if we choose $q = 1$ for the second term of (3), it is the $\ell_{1,1}$ norm usually used for earlier frame and wavelet based image restoration models, which is known as the anisotropic ℓ_1 norm of the frame coefficients. When we choose $q = 2$, the second term of (3) is called the isotropic ℓ_1 norm, which can be understood as an isotropic ℓ_1 of the frame coefficients using an ℓ_2 norm to combine different frame bands at a given location and decomposition level. The isotropic ℓ_1 norm proposed in [30] was shown to be superior than the anisotropic ℓ_1 norm, and therefore we select $q = 2$ to make a comparison of the two algorithms in the section of numerical experiments in this paper.

B. Review of the Split Bregman Algorithm for (3)

In this subsection, we mainly review the classical split bregman [30] for the minimization problem (3) or equivalently (5).

The split bregman algorithm was first proposed in [41] and was shown to be powerful in [41] and [42] when it is applied to solving various PDE based image restoration problems, e.g., ROF and nonlocal variational models. The initial idea of operator splitting, i.e. the transformation from (3) to (5) was early proposed in our previous work [55]. The split bregman algorithm has been recently proven to be equivalent to the alternating direction method of multipliers (ADMM) [43]–[45] applied to the augmented Lagrangian [46]–[48] of the problem (3):

$$\begin{aligned}\mathcal{L}(u, \alpha, v) = & \frac{1}{2} \|Au - f\|_2^2 + \|\lambda \cdot \alpha\|_{1,q} \\ & + \langle v, Wu - \alpha \rangle + \frac{\mu}{2} \|Wu - \alpha\|_2^2.\end{aligned}\quad (6)$$

The corresponding augmented Lagrangian(AL) method can be described as the following two alternative minimization steps:

$$\begin{cases} (u^{k+1}, \alpha^{k+1}) = \arg \min_{u, \alpha} \mathcal{L}(u, \alpha, v^k) & (\text{step a}) \\ v^{k+1} = v^k + \delta(Wu^{k+1} - \alpha^{k+1}) & (\text{step b}) \end{cases}\quad (7)$$

for some $\delta > 0$. And one can approximates the solutions (u^{k+1}, α^{k+1}) of step a of (7) by the following alternative two steps:

$$\begin{cases} u^{k+1} = \arg \min_u \mathcal{L}(u, \alpha^k, v^k) \\ \alpha^{k+1} = \arg \min \mathcal{L}(u^{k+1}, \alpha, v^{k+1}) \end{cases}$$

This way we get the split bregman algorithm for the optimization problem (5). Choosing $\delta = \mu$ (see [6], [49], [50]),

Algorithm 1 The Split Bregman Algorithm [2], [40], [29]

Given an observed image f , initialize $\alpha^0 = v^0 = 0$.

While stopping criterion is not met **Do**

1. update u :
 $u^{k+1} = (A^T A + \mu I)^{-1}(A^T + \mu W^T(\alpha^k - v^k))$
 2. update α :
 $\alpha^{k+1} = \tau_{\lambda/\mu}^q(Wu^{k+1} + v^k)$
 3. update v :
 $v^{k+1} = v^k + Wu^{k+1} - \alpha^{k+1}$
 4. $k = k + 1$.
- end while**
-

we have the split bregman method in Algorithm 1, based on some computations of the augment Lagrangian $\mathcal{L}(u, \alpha, v)$.

The minimization with respect to u is a least square problem with the normal equation

$$(A^T A + \mu I)u^{k+1} = A^T f + \mu W^T(\alpha^k - v^k) \quad (8)$$

where A is a projection operator for image inpainting problems, the coefficient matrix $(A^T A + \mu I)$ is a diagonal matrix, and therefore this inverse problem can be solved efficiently.

The second subproblem of Algorithm 1 can also be solved rather efficiently. For the case $q = 1$, α^{k+1} can be obtained by soft-thresholding as follows (see [51], [52]):

$$\alpha^{k+1} = \tau_{\lambda/\mu}^1(Wu^{k+1} + v^k) \quad (9)$$

where the soft-thresholding operator $\tau_\tau^1(\xi) := \frac{\xi}{|\xi|} \max\{|\xi| - \tau, 0\}$ (note that this operator is an entry-wise operation). The superscript 1 of $\tau_{\lambda/\mu}^1$ indicates that it is the shrinkage operator corresponding to the $\ell_{1,1}$ -norm, and we call this case as the anisotropic shrinkage. For $q = 2$, it becomes a more complicated situation, and we will need some proper assumptions on the parameter λ in order to obtain neat analytical formula for α^{k+1} . Define J_l , for each $0 \leq l \leq L-1$, to be a subset of the index set of \mathcal{I} , and denote $J_l^c := \mathcal{I} \setminus J_l$. Assume that

$$\lambda_{l,j} = \begin{cases} 0, & j \in J_l \\ \tau_l, & j \in J_l^c \end{cases} \quad (10)$$

where for each $0 \leq l \leq L-1$, $\tau_l \in \mathbb{R}^n$ denotes a vector. Due to the structure of the framelet decomposition Wu , the bands $(0, 0) \in \mathcal{I}$ of Wu corresponds to the low frequency components of u , which is not sparse in general. So we should not penalize the ℓ_1 -norm of it (see [30] in details). Under this assumption about λ , we obtain

$$\alpha^* = \arg \min_{\alpha} \|\lambda \cdot \alpha\|_2 + \frac{\mu}{2} \|\alpha - \xi\|_2^2 = \tau_{\lambda/\mu}^2(\xi) \quad (11)$$

Here we have

$$(\tau_{\lambda/\mu}^2(\xi))_{l,j} = \begin{cases} \xi_{l,j}, & j \in J_l \\ \frac{\xi_{l,j}}{F_l} \max\{F_l - \tau_l/\mu, 0\}, & j \in J_l^c \end{cases} \quad (12)$$

where $F_l = (\sum_{j \in J_l^c} |\xi_{l,j}|^2)^{\frac{1}{2}}$. Similarly, the superscript 2 of τ_τ^2 indicates that it is the shrinkage operator corresponding to the $\ell_{1,2}$ -norm, and we call τ_τ^2 as the isotropic shrinkage. The

convergency analysis of the split bregman to solve the analysis based approach (3) was given in [2].

Remarks for Algorithm 1: For the case $q = 2$ and λ not satisfying the assumption (10), one can still obtain an analytic form of the solution at least theoretically, for the optimization problem (11). However, the analytic form is rather complicated. So, in this paper, we will choose the λ that satisfies the condition of (10), and we will not go into details of solving (11) for general λ . Since the proof of convergence of the split bregman algorithm given by [2] is meant for general case, it directly indicates the convergence of the Algorithm 1.

III. OUR PROPOSED MODEL AND ALGORITHM

A. The Weighted ASBA

The original model (3) is based on the ℓ_1 norm, which is a popular sparsity enforcement regularization due to its convexity. However, in general, the non-convex sparse regularization prefers an even more sparse solution and usually has a better theoretical property, such as the widely used ℓ_p norm ($0 \leq p < 1$) models [7], [8], [11], [65]. While there have existed many works based on non-convex penalization for image processing, there have existed only few specific works for image inpainting based on wavelet frame, to our best knowledge. The major difficulties with the existing nonconvex algorithms are that the global optimal solution cannot be efficiently computed for now, the behavior of a local solution is also hard to analyze [58] and more seriously the prior structural information of the solution is hard to be incorporated. In this paper, we present a non-convex sparse model based on the convex model (3) and proposed a multi-stage convex relaxation scheme for solving it via our proposed iterative support detection [13], which is able to incorporate the structural information of the underlying solution and has proved to be better than the pure ℓ_1 solution in theory in many cases [13]. We will also demonstrate its advantages over some non-convex sparse regularization like ℓ_p norm ($0 \leq p < 1$) based models, which also fail to make use of the structural information of the wavelet frame coefficients.

We first rewrite the last term of (3) similar to the form of variational methods as follows:

$$(ASBA) \quad \min_u \frac{1}{2} \|Au - f\|_2^2 + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n \|\lambda_{l,i} \circ W_{l,i} u\|_q \right) \quad (13)$$

where $W_{l,i} u$, and $\lambda_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ are vectors, which are the sets of $W_{l,i,j} u$, $\lambda_{l,i,j}$, $j \in \mathcal{I}$, respectively. n equals to the number of total pixels, and $|\mathcal{I}|$ denotes the total number of bands in given level at each pixel, e.g., $|\mathcal{I}| = 9$ for the piecewise linear B-spline framelets. In order to better illustrate these notations, we use the total variation as an example. As we know that the total variation is summation of variation over all pixels, for each given pixel, the variation has two directions, which are horizontal and vertical ones, respectively. Generalizing this to the case of wavelet framelets, we can similarly think that the $W_{l,i}$ has $|\mathcal{I}|$ directions in each level at given pixel. Here,

we use the following vector computation rule:

$$\begin{aligned} a \circ b &= (a_1 \times b_1, a_2 \times b_2, \dots, a_p \times b_p), \\ a \circ b \circ c &= (a_1 \times b_1 \times c_1, a_2 \times b_2 \times c_2, \dots, a_p \times b_p \times c_p) \end{aligned}$$

where vector $a = (a_1, a_2, \dots, a_p)$, $b = (b_1, b_2, \dots, b_p)$, and $c = (c_1, c_2, \dots, c_p)$ have the same dimensionality. Note that the output of this vector computation rule, is not a scalar, but a vector of the original dimensionality. For the complicated case $q = 2$, we choose the $\lambda_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ in line with (10):

$$\lambda_{l,i,j} = \begin{cases} 0, & j \in J_l \\ \varrho_l, & j \in J_l^c \end{cases} \quad (14)$$

where for each $0 \leq l \leq L-1$, $i \in \{1, 2, \dots, n\}$, and ϱ_l is a scalar. Namely, in level l at given pixel i , when $j \in J_l$, the correspondent scalar value $\lambda_{l,i,j}$ equals to 0. We should point out that the necessity to introduce the equivalent form of the model ASBA was just for the convenient introduction for our new proposed model, and the better comparison between these two models.

Now we introduce our new model and corresponding algorithm. The new model is named as Weighted ASBA and has the form as follows:

$$\min_u \frac{1}{2} \|Au - f\|_2^2 + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n \|T_{l,i} \circ (\lambda_{l,i} \circ W_{l,i} u)\|_q \right) \quad (15)$$

where $T_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ are the weights, which are related with u . The only difference between the old model (13) and our proposed new model, is the introduction of the weights $T_{l,i}$. With regard to the old model ASBA, traditionally, we just solved a plain convex optimization minimization problem one time and get the final result, and it has been generally accepted that the sparse penalization is uniform, i.e. the setting of weights $T_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ are always ones.

While it is well known that by choosing proper weights $T_{l,i}$, the weighted model (15) should perform much better than the plain ℓ_1 model (13), how to choose the weights is not trivial and even very difficult. One difficulty is that we do not have a lot of information about the true image \bar{u} beforehand and the other is that once we have some information about the true solution, we need to find a good way to make use of it to calculate the weights $T_{l,i}$. In this paper, we proposed to make use of the idea of the iterative support detection [13] to extract the reliable information about the true solution and set up correspondingly weights. The corresponding new algorithm is an alternating optimization procedure, which repeatedly applies the following two steps:

- Step 1: First we optimize u with $T_{l,i}$ fixed (initially $\vec{1}$): this is a convex problem in u .
- Step 2: Second we determine the value of $T_{l,i}$ according to the current u . The value of $T_{l,i}$ will be used in the Step 1 of the next iteration.

In order to achieve more direct comparison between the two different kinds of algorithms, we show the flow chart of SB method and our proposed method in Figure 1, we can see that for our proposed method, the final result is obtained from multi-stage process by solving a series of weighted ASBA problems. The scheme of alternating optimization help

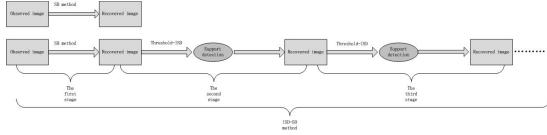


Fig. 1. The algorithm framework of single stage SB method (the first row) and our proposed multi-stage method ((JL)ISD-SB, see the second row).

decouple the original too complicated model (15) into two reasonably easier sub-problems. We start from solving the plain ℓ_1 model (13), then improve the solution gradually via the above multi-stage procedure. At each stage (iteration) in the process, the weights will change according to the most recently recovered image via adaptive learning. This full procedure and Step 2 will be introduced in more details in next section, where the iterative support detection is reviewed.

B. Step 1: Solving a Weighted ASBA Given Weights

Now we come to how to solve the weighted ASBA once the weights are given, i.e. how to realize Step 1. We transform the problem (15) to

$$\begin{aligned} \min_{\alpha, u} \quad & \frac{1}{2} \|Au - f\|_2^2 + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n \|T_{l,i} \circ (\lambda_{l,i} \circ \alpha_{l,i})\|_q \right), \\ \text{s.t.} \quad & \alpha_{l,i} = W_{l,i} u. \end{aligned} \quad (16)$$

Then the augment Lagrangian function of (15) can be written as

$$\begin{aligned} \mathcal{L}(\alpha, u) = & \frac{1}{2} \|Au - f\|_2^2 \\ & + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n (\|T_{l,i} \circ (\lambda_{l,i} \circ \alpha_{l,i})\|_q) \right) \\ & + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n (\langle v_{l,i}, W_{l,i} u - \alpha_{l,i} \rangle) \right) \\ & + \sum_{l=0}^{L-1} \left(\sum_{i=1}^n \left(\frac{\mu}{2} \|W_{l,i} u - \alpha_{l,i}\|_2^2 \right) \right). \end{aligned}$$

We note that α is the complete set of $\alpha_{l,i,j}$, where $l \in 0 \leq l \leq L-1$, $i \in \{1, 2, \dots, n\}$, $j \in \mathcal{I}$. The original split bregman algorithm can be applied here with some modifications due to the existence of $T_{l,i}$ as follows.

This joint minimization problem of (α, u) can be carried in parallel. Specifically, the minimization with respect to u is attained by

$$(A^T A + \mu I)u = A^T f + \mu W^T(\alpha - v/\mu) \quad (17)$$

Note that similar to α , v is the complete set of $v_{l,i,j}$, where $l \in 0 \leq l \leq L-1$, $i \in \{1, 2, \dots, n\}$, $j \in \mathcal{I}$. The optimization problem with respect to $\alpha_{l,i}$ is obtained by

$$\begin{aligned} \min_{\alpha_{l,i}} \quad & \|T_{l,i} \circ (\lambda_{l,i} \circ \alpha_{l,i})\|_q + \langle v_{l,i}, W_{l,i} u - \alpha_{l,i} \rangle \\ & + \frac{\mu}{2} \|W_{l,i} u - \alpha_{l,i}\|_2^2 \end{aligned} \quad (18)$$

Algorithm 2 Solving the Weighted ASBA (15) Given Weights

Given an observed image f , and the projection operator A . the outsider iteration number is s , the outcome u of the $(s-1)$ th alternative optimization, initialize $\alpha^0 = v^0 = W u^{(s-1)}$.

While stopping criterion is not met **Do**

1. update u :

$$u^{k+1} = (A^T A + \mu I)^{-1}(A^T + \mu W^T(\alpha^k - v^k))$$
2. update α :
 for the case $q = 1$:

$$\alpha_{l,i}^{k+1} = \max(|\xi_{l,i}^k| - (T_{l,i} \circ \lambda_{l,i})/\mu, 0) \cdot \frac{\xi_{l,i}^k}{|\xi_{l,i}^k|}$$

 where $\xi_{l,i}^k = W_{l,i} u^{k+1} + v_{l,i}^k$.
 for the case $q = 2$:

$$\text{if } j \in J_l, \alpha_{l,i,j}^{k+1} = W_{l,i,j} u^{k+1} + \frac{v_{l,i,j}^k}{\mu}.$$

$$\text{if } j \in J_l^c,$$

$$\alpha_{l,i,j}^{k+1} = \max(\|T_{l,i} \circ (\xi_{l,i}^k)'\|_2 - (T_{l,i,j} \cdot \lambda_{l,i,j})/\mu, 0)$$

$$\cdot \frac{\xi_{l,i,j}^k}{\|T_{l,i} \circ (\xi_{l,i}^k)'\|_2}, \text{ where } \xi_{l,i,j}^k = W_{l,i,j} u^{k+1} + v_{l,i,j}^k,$$

$$(\xi_{l,i}^k)' = W_{l,i}' u^{k+1} + (v_{l,i}^k)'$$
.
3. update v :

$$v^{k+1} = v^k + W u^{k+1} - \alpha^{k+1}$$
4. $k = k + 1$.

end while

And the above formula is equivalent to

$$\min_{\alpha_{l,i}} \|T_{l,i} \circ (\lambda_{l,i} \circ \alpha_{l,i})\|_q + \frac{\mu}{2} \|\alpha_{l,i} - W_{l,i} u - \frac{v_{l,i}}{\mu}\|_2^2 \quad (19)$$

For the case of $q = 1$, the solution of subproblem (19) can be obtained by the well known soft-shrinkage operator:

$$\alpha_{l,i} = \max(|\zeta_{l,i}| - (T_{l,i} \circ \lambda_{l,i})/\mu, 0) \cdot \frac{\zeta_{l,i}}{|\zeta_{l,i}|} \quad (20)$$

where $\zeta_{l,i} = W_{l,i} u + \frac{v_{l,i}}{\mu}$, $0/0 = 1$, and $0 \cdot (0/0) = 0$ is assumed. For the case of $q = 2$, it becomes more complicated since the elements in the weights vector $T_{l,i} \in \mathbb{R}^{|\mathcal{I}|}$ are the mixture of 0 and 1. In addition, note that if $j \in J_l$, $\lambda_{l,i,j} = 0$. For the paper conciseness, we will omit the derivation and only give the outcome here. If $j \in J_l$, we get

$$\alpha_{l,i,j} = W_{l,i,j} u + \frac{v_{l,i,j}}{\mu} \quad (21)$$

For the cases where $j \in J_l^c$, we let $\alpha'_{l,i}$ denote a vector, whose elements are $\alpha_{l,i,j}$. Clearly, $\alpha'_{l,i}$ is the subvector of $\alpha_{l,i}$. In what follows, the similar form to $\alpha'_{l,i}$ has the same meaning. Then we can obtain the complementary solution of $\alpha_{l,i}$. If $j \in J_l^c$, we obtain

$$\begin{aligned} \alpha_{l,i,j} = & \max(\|T_{l,i} \circ \xi'_{l,i}\|_2 - (T_{l,i,j} \cdot \lambda_{l,i,j})/\mu, 0) \\ & \cdot \frac{\xi'_{l,i}}{\|T_{l,i} \circ \xi'_{l,i}\|_2} \end{aligned} \quad (22)$$

where $\xi_{l,i,j} = W_{l,i,j} u + \frac{v_{l,i,j}}{\mu}$, $\xi'_{l,i} = W'_{l,i} u + \frac{v'_{l,i}}{\mu}$, $0/0 = 1$, and $0 \cdot (0/0) = 0$ is assumed in the same. The procedure is summarized in Algorithm 2.

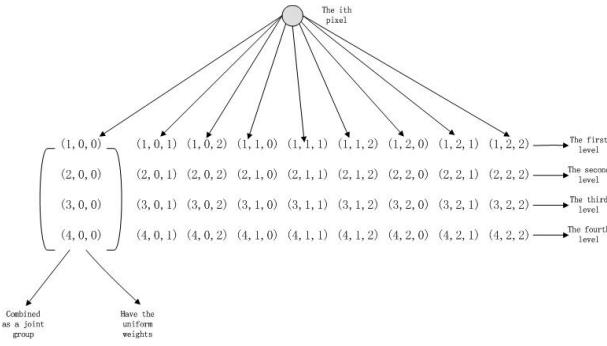


Fig. 2. This is to show the joint level support detection. Specifically, we combine the given frame bands of all levels at the given pixel as a vector (joint group), and calculate its ℓ_2 norm magnitude, and all its components will be set to have the same weights (either 0 or 1).

Notice that the above procedure is in fact a selective shrinkage procedure. The original ℓ_1 model obtains a sparse solution via the soft shrinkage operator. However, this kind of shrinkage has a fatal disadvantage, i.e., it shrinks the true nonzero components as well, and reduces the sharpness of the recovered images. By using 0–1 weights, some components will not be shrunk if we believe that they are unlikely to be zero. In such cases, their corresponding weights are set as 0. This is a natural setting and the advantages of the 0–1 weights over other kinds of weights have been demonstrated in either theoretic or practical point of view in [13] and will be reviewed in Part C.

Remarks for Algorithm 2: note that initializing $\alpha^0 = v^0 = Wu^{(s-1)}$ in Algorithm 2 is actually a warm-starting in the whole algorithmic framework based on alternative optimization. Since the proof of convergence of the split bregman algorithm given by [2] is very general, it directly indicates the convergence of the Algorithm 2.

C. Step 2: Weights Determination Based on Iterative Support Detection

In this part, we explain the way of determining $T_{l,i}$ of Step 1. While the idea is originally coming from our proposed iterative support detection (ISD) in pure sparse signal recovery [13], we propose a specific implementation to make use of the multi-level structures of the coefficients of wavelet frame, when considering the image inpainting problem.

To make the paper self-contained, we first briefly review the idea of ISD. While the idea of exploiting available partial support detection have arisen in several subsequent literatures, see [14]–[19], ISD also focus more on how to extract this partial support information instead. Compressive sensing [21], [22] reconstructs a unknown sparse signal from a small set of linear projections. Let \bar{x} denote a k -sparse signal, and let $b = Ax$ represent a set of m linear projections of \bar{x} . The general reconstruction optimization problem is Basis Pursuit (BP) problem

$$(BP) \quad \min_x \|x\|_1 \quad s.t. \quad Ax = b. \quad (23)$$

However, unlike the BP problem which is a one-stage convex relaxation method, ISD is a multi-stage convex relaxation

Algorithm 3 The (JL)ISD-SB Algorithm

- Given an observed image f and the projection operator A .
1. Set the iteration $s \leftarrow 0$ and initialize the set of detected entries $I^{(s)} \leftarrow \emptyset$
 2. while the stopping condition is not satisfied,
 - (a) $T^{(s)} \leftarrow (I^{(s)})^C := \Omega \setminus I^{(s)}$;
 - (b) $u^s \leftarrow$ solve weighted ASBA (15) for $T = T^{(s)}$;
(Step 2: using Algorithm 2)
 - (c) $I^{(s+1)} \leftarrow$ support detection using $u^{(s)}$ as the reference;
(Step 1:)
 - (d) $s \leftarrow s + 1$.
-



Fig. 3. 6 test images (from the left to right, up to down: checkerboard, lena, cameraman, boat, bowl, and man, respectively.). Sizes of them are: 200 × 200, 240 × 232, 256 × 256, 256 × 256, 256 × 256, 256 × 256, respectively.

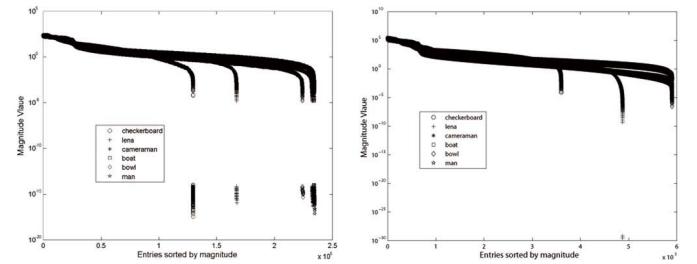


Fig. 4. The sorted magnitudes of the wavelet frame coefficients of the 6 test images (first column), the sorted joint-level group magnitudes of the wavelet frame coefficients of the 6 test images (second column), both of them have the property of fast decaying property.

method, alternatively calling its two components: support detection and signal reconstruction. ISD starts from solving a plain BP problem. If the BP model returns a correct sparse solution, things are fine and ISD stops there. Otherwise, from the incorrect reconstruction, support detection will be performed to identify an index set I , which contain some elements of $\text{supp}(\bar{x}) = \{i : \bar{x}_i \neq 0\}$. After the acquiring of support detection, the corresponding elements will be truncated from the ℓ_1 regularization term and the resulted model is as follows

$$(\text{Truncated BP}) \quad \min_x \|x_T\|_1 \quad s.t. \quad Ax = b \quad (24)$$

where $T = I^C$ and $\|x_T\|_1 = \sum_{i \notin I} |x_i|$. If the support detection set $I = \text{supp}(\bar{x})$, the solution of (24) is, of course,

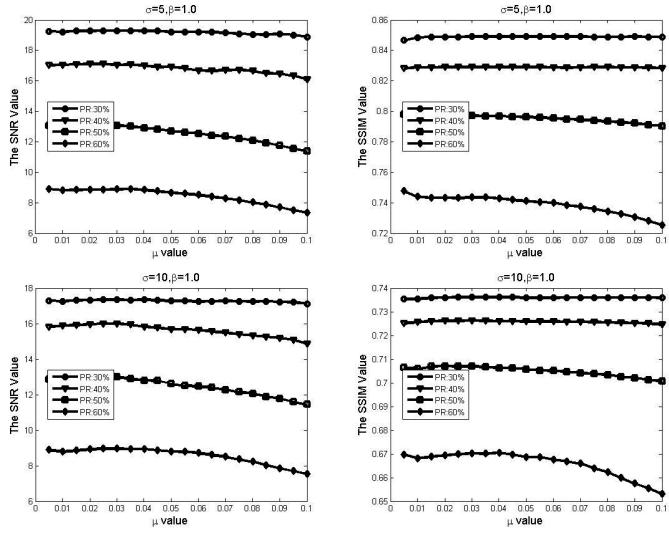


Fig. 5. The SNR and SSIM value trend of ISD-SB method as the parameter μ varies, where $\rho = 3$, the maximum stage number is set as 4, and the test image is lena.

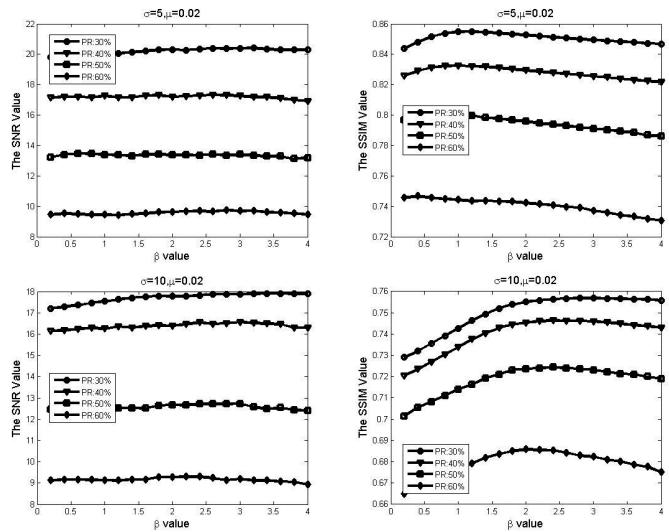


Fig. 6. The SNR and SSIM value trend of ISD-SB method as the parameter β varies, where $\rho = 3$, the maximum stage number is set as 4, and the test image is lena.

equal to the true original signal \bar{x} . Even if I does not have enough of $\text{supp}(\bar{x})$ for an exact reconstruction, those entries of $\text{supp}(\bar{x})$ in I will help (24) reconstruct a better solution compared to (23). From this better solution, support detection will be able to identify more entries of $\text{supp}(\bar{x})$ and then yield an even better I . With this method, the two components of ISD work together to gradually recover $\text{supp}(\bar{x})$ and improve the reconstruction result.

In most cases, even if the true signal \bar{x} itself is not sparse or compressible, its representation under a certain basis, frame, or dictionary that is sparse or compressible. In such a case, $\bar{y} = W\bar{x}$ is sparse or compressible for a certain linear transform W . So, instead of minimizing $\|x\|_1$ and $\|x\|_T$, one should minimize $\|Wx\|_1$ and $\|(Wx)_T\|_1$ as a substitute respectively.

TABLE I
THE COMPARISON OF SNR AND SSIM VALUES BETWEEN SB, MDAL,
ISD-SB, JLISD-SB, IISD-SB METHOD. NOISE LEVEL: $\sigma = 5$

PR	Image name	SB		MDAL		ISD-SB		JLISD-SB		IISD-SB	
		SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM
30%	checkerboard	19.170/0.8922	27.750/0.9164	28.320/0.9177	25.910/0.9098	30.160/0.9210					
	lena	10.470/0.8191	20.290/0.8389	19.20/0.8495	19.63/0.8497	27.09/0.8814					
	cameraman	16.790/0.5533	17.990/0.5418	18.890/0.5697	19.410/0.5632	23.920/0.6000					
	boat	13.890/0.7544	14.550/0.7631	15.700/0.7821	15.970/0.7827	20.590/0.8357					
	bowl	14.520/0.5882	15.420/0.6053	16.190/0.6217	16.390/0.6243	19.10/0.6625					
40%	checkerboard	17.11/0.8734	26.550/0.9126	26.770/0.9152	23.10/0.9048	29.82/0.9199					
	lena	8.570/0.7755	18.39/0.8137	17.120/0.8263	17.780/0.8273	26.820/0.8752					
	cameraman	14.590/0.5310	16.690/0.5525	17.430/0.5525	18.070/0.5445	23.780/0.5967					
	boat	12.630/0.7202	13.120/0.7207	14.080/0.7527	14.650/0.7504	20.31/0.8324					
	bowl	13.370/0.5513	14.130/0.5750	14.920/0.5955	15.040/0.6005	19.160/0.6639					
50%	checkerboard	14.950/0.8431	25.130/0.9057	25.330/0.9140	21.84/0.9019	29.44/0.9189					
	lena	7.230/0.7305	16.670/0.7840	13.01/0.7981	15.36/0.8015	26.680/0.8647					
	cameraman	14.190/0.4996	14.610/0.4841	15.970/0.5281	16.250/0.5156	23.560/0.5928					
	boat	11.490/0.6738	11.570/0.6684	12.620/0.7098	12.970/0.7159	19.990/0.8292					
	bowl	12.400/0.5107	12.950/0.5314	13.620/0.5585	13.820/0.5715	19.200/0.6674					
60%	checkerboard	12.580/0.7892	23.39/0.8973	20.70/0.9018	18.47/0.8791	29.16/0.9192					
	lena	8.540/0.6657	14.660/0.7407	9.330/0.7500	11.860/0.7611	26.180/0.8546					
	cameraman	12.730/0.4606	13.470/0.4479	14.060/0.4939	14.690/0.4814	23.200/0.5849					
	boat	10.36/0.6625	10.19/0.5988	11.28/0.6814	11.67/0.6807	19.38/0.8230					
	bowl	11.550/0.4627	11.700/0.4693	12.330/0.5076	12.510/0.5184	18.860/0.6676					

TABLE II
THE COMPARISON OF SNR AND SSIM VALUES BETWEEN SB , MDAL,
ISD-SB, JLISD-SB, IISD-SB METHOD. NOISE LEVEL: $\sigma = 10$

PR	Image name	SB		MDAL		ISD-SB		JLISD-SB		IISD-SB	
		SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM	SNR/SSIM
30%	checkerboard	18.140/0.8508	23.850/0.8730	24.110/0.8775	22.400/0.8649	25.610/0.8877					
	lena	10.020/0.7096	17.980/0.7304	17.560/0.7413	18.100/0.7470	22.200/0.7712					
	cameraman	14.940/0.4151	15.490/0.4172	15.950/0.4249	16.370/0.4326	19.050/0.4684					
	boat	12.060/0.6149	12.330/0.6188	12.870/0.6297	13.370/0.6499	15.680/0.7078					
	bowl	11.420/0.4689	11.670/0.4748	11.810/0.4845	12.230/0.4904	14.310/0.5599					
40%	checkerboard	16.08/0.8315	23.24/0.8742	23.74/0.8783	21.44/0.8615	25.57/0.8862					
	lena	8.470/0.6803	16.970/0.7224	15.090/0.7260	16.580/0.7365	22.05/0.7763					
	cameraman	14.040/0.4008	14.720/0.4009	15.220/0.4143	16.060/0.4246	18.630/0.4635					
	boat	11.370/0.5937	11.500/0.5895	12.180/0.6119	12.740/0.6333	15.590/0.7029					
	bowl	11.130/0.4431	11.40/0.4511	11.530/0.4660	11.910/0.4675	13.910/0.5565					
50%	checkerboard	14.250/0.8053	22.28/0.8730	21.42/0.8766	19.650/0.8563	25.340/0.8820					
	lena	6.990/0.6417	15.60/0.7000	12.40/0.7053	14.580/0.7188	21.70/0.7745					
	cameraman	13.340/0.3822	13.680/0.3794	14.530/0.3983	14.980/0.4086	18.280/0.4628					
	boat	10.590/0.5618	10.570/0.5545	11.330/0.5817	11.860/0.6027	15.300/0.6984					
	bowl	10.750/0.4118	10.70/0.4171	11.070/0.4282	11.400/0.4352	13.410/0.5466					
60%	checkerboard	11.970/0.7519	21.26/0.8665	18.860/0.8677	17.980/0.8433	25.140/0.8815					
	lena	5.690/0.5888	13.61/0.6751	8.590/0.6599	12.080/0.7001	21.380/0.7682					
	cameraman	12.400/0.5537	12.680/0.5476	13.430/0.5721	14.110/0.5913	17.900/0.6469					
	boat	9.580/0.5165	9.580/0.4995	10.150/0.5361	10.740/0.5641	15.030/0.6887					
	bowl	10.40/0.3761	10.34/0.3727	10.620/0.3882	10.80/0.3935	12.990/0.5384					
60%	checkerboard	10.98/0.5812	10.86/0.5622	11.71/0.6054	12.30/0.6330	16.90/0.7664					

From the inexact intermediate reconstruction, ISD tries to obtain a reliable support detection, which can be able to take advantages of the features and prior information about the true signal \bar{x} . The authors in [13] focus on the sparse or compressible signals with components having a fast decaying distribution of nonzeros. For these kinds of signals, they proposed to perform the support detection by thresholding the solution of (24), and called the corresponding ISD algorithm *threshold-ISD*. Notice that the fast decaying property is a mild assumption because in practice most natural images satisfy this property in a appropriate basis, for example, wavelets, curvelets and et al.

As for the image inpainting problem, the above support detection approach *threshold-ISD* can be applied in a similar way as follows. We present effective detection strategies for images with a fast decaying distribution of coefficients under wavelet frame transforms (we call them fast decaying images here and following in brief). Our methods are based on thresholding and for s -th stage, we have

$$I^{(s+1)} := \{(l, i, j) : |W_{l,i,j} u^{(s)}| > \epsilon^{(s)}\}, \quad (25)$$

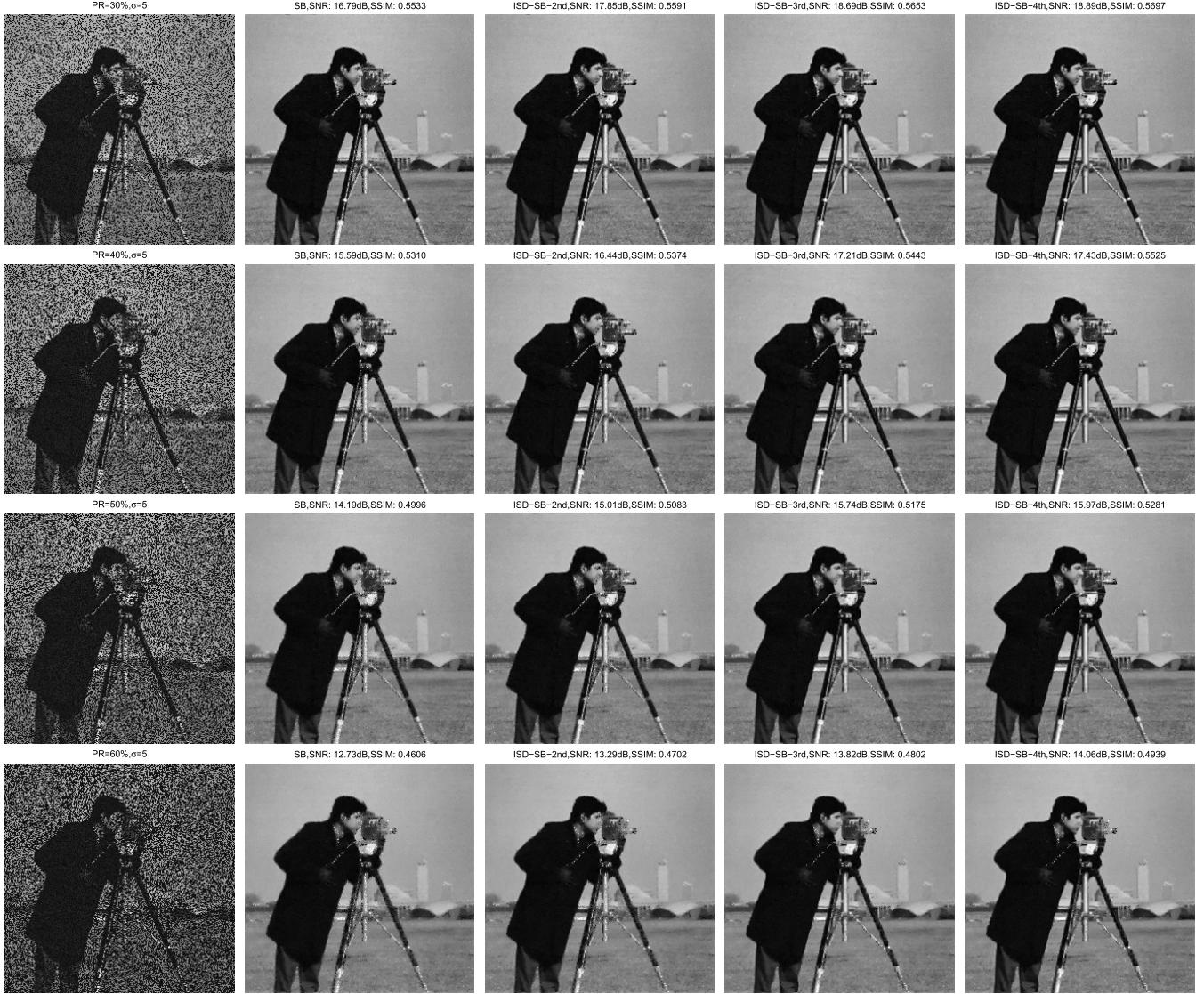


Fig. 7. The intermediate stage results of ISD-SB method are showed, where the second, third, fourth and fifth columns are corresponding to the results of the first stage, second stage and third stage and fourth stage, respectively. The first row, the second row, the third row and the fourth row correspond to the results of different PR values from 30%, 40%, 50%, to 60%, respectively. Noise level: $\sigma = 5$. Test image: cameraman.

$s = 0, 1, 2, \dots$, The weights $T_{l,i,j}^{(s)}$ equals 0 if the corresponding element belongs to the support detection $I^{(s)}$, or 1 otherwise. Before discussing the choice of $\epsilon^{(s)}$, we point out that support detection sets $I^{(s)}$ are not necessarily increasing and nested, i.e., $I^{(s)} \subset I^{(s+1)}$ may not hold for all s . This is very important because the support detection we get from the currently solution may contains wrong detections by $\epsilon^{(s)}$ thresholding. Not requiring $I^{(s)}$ to be monotonic leaves the chance for support detection to remove previous wrong detections, and this makes $I^{(s)}$ less sensitive to $\epsilon^{(s)}$, thus making $\epsilon^{(s)}$ easier to choose. In [13], we have proved that ISD can tolerate certain ratio of wrong detections and still achieve a better recovery.

As for the threshold rule, we set

$$\epsilon^{(s)} := \max\{|W_{l,i,j} u^{(s)}|\} / \rho^{(s+1)} \quad (26)$$

with $\rho > 0$. The above is only a plain extension of the

original ISD to our problem and we can make the support detection even more effective by considering the structures of the wavelet coefficients. It is known that the edges of an image should correspond to large coefficients consistently for all levels, in wavelet frames. Considering this specific characteristic, we propose a corresponding effective support detection strategy, which is shown in Figure 2. Specifically, for a given pixel, we combine the given frame bands of all levels as a joint group, the magnitudes of these groups are measured as $\|\cdot\|_2^2$, and the weights in each group are uniform (0 or 1), i.e., in given bands j at given pixel i , the weights in each level ($0 \leq l \leq L - 1$) are same. In order to better illustrate this strategy, we introduce a 2-mode tensor, denoted as Θ , whose elements are the aforementioned magnitudes of these groups taken over all pixels, denoted as $\Theta_{i,j}$, where the subscript i denotes the i th pixel, and j denotes the specific frame bands of a given pixel. The cardinality of Θ is $n \times |\mathcal{I}|$,

where n is equal to the number of total pixels, and $|\mathcal{I}|$ is the total number of bands in given level at each pixel, e.g., $|\mathcal{I}| = 9$ if we use the piecewise linear B-spline framelets in this paper. It should be pointed out that the sorted magnitudes of tensor Θ still have the property of fast decaying, and the support detection method is based on thresholding:

$$I^{(s+1)} := \{(i, j) : |\Theta_{i,j}| > \varepsilon^{(s)}\}, \quad (27)$$

$s = 0, 1, 2, \dots$. Similarly, we set the threshold value:

$$\varepsilon^{(s)} := \max\{|\Theta_{i,j}|\}/\eta^{(s+1)}. \quad (28)$$

We named this support detection method as joint level iterative support detection method (JLISD method for short).

An excessively large ρ and η results in too many false detections and lead to low solution quality, while an excessively small ρ and η tends to cause a large number of iterations. ISD will be quite effective with an appropriate ρ and η , though the proper range of ρ and η might be case-dependent, due to the No Free Lunch theorem [59]. However, numerical experiments have shown that the performance of ISD is not very sensitive to the choice of ρ and η .

Now we would like to further explain the features of ISD, as a specific 0–1 weighting scheme, in an intuitive way, compared to other similar multi-stage algorithms including the well known iterative reweighted ℓ_1 algorithm [60]. Firstly, we emphasize the importance of the partial support information as prior to better reconstruct the original sparse signals (images), and inspire ones to develop new and effective implementations in their own particular problems, for example, edge detection in the total variation regularized image reconstruction [20]. Secondly, its performance over the single stage pure ℓ_1 model has been proved rigorously in [13] once we can detect correct partial support information (by thresholding, for example, in this paper), while most of the related alternatives such as the reweighted ℓ_1 algorithm [60] do not have such rigorous theoretical guarantees. Thirdly, for the reweighted ℓ_1 algorithm, the weights are usually set like $\theta_i = \frac{1}{|x_i| + \epsilon}$. In [13], we have pointed out that the tuning parameter $\epsilon > 0$ is the key parameter and should be determined carefully. Roughly, ϵ should not be a fixed value and should decrease from a large value to a small value. In an extreme case, if ϵ is always fixed to 0, then we would not get a better solution for the next stage, even there is no numerical trouble of dividing 0, because we just passively make use of the all the information of the current solution without any filtering out the inaccurate information. From our analysis, the choice of ϵ is like controlling the extraction of the useful information and suppressing the distortion of the recovery noise. From the point of view, our 0–1 weights of ISD is a more explicit and straightforward way to imply the idea of making use of the correct information (mainly about the locations of components of large magnitude and setting the corresponding weights as 0) and give up the rest too noisy information (for those components of small magnitudes, they are mostly overwhelmed by the recovery noise and therefore there is very little meaning to set different weights according to their magnitudes. So it is more reasonable to set the same weights as 1).



Fig. 8. The intermediate stage results of JLISD-SB method are showed, where the first, second and the third columns are corresponding to the results of the first stage, second stage and third stage, respectively. The first row, the second row, the third row and the fourth row correspond to the results of different PR values from 30%, 40%, 50%, to 60%, respectively. Noise level: $\sigma = 5$. Test image: cameraman.

D. The Algorithm Framework for New Model (15)

Now we summarize the algorithmic framework of the multi-stage convex relaxation for the weighted model (15) based on ISD. As mentioned above, the algorithm repeatedly performs two steps: support detection to determine the 0–1 weights and solving the resulted weighted ℓ_1 model using the splitting bregman (SB, for short) algorithm. Therefore, we named our new algorithm as ISD-SB when the support detection is based on (25), JLISD-SB when the support detection is based on (27).

Ω denotes the universal set of $(l, i, j), 1 \leq l \leq L - 1, i \in \{1, 2, \dots, n\}, j \in \mathcal{I}$.

Remarks for Algorithm 3: note that (15) in step 2(b) reduces to ASBA in iteration 0, where $I^{(0)} = \emptyset$, and $T^{(0)} = \Omega$. The weighted ASBA (15) in 2(b) of Algorithm 3 can be solved by Algorithm 2. The support detection of in 2(c) has been introduced in the Section III(C).

The computational time complexity of ISD-SB and JLISD-SB is not necessarily several times more than that of SB if the warm starting (i.e. the output of the current stage (outside iteration) is used to be the input of the next stage.) and looser stopping tolerance (except the final stage) are adopted, as mentioned in the original ISD paper, where their computational time can be even almost equal sometimes.



Fig. 9. The visual comparisons between SB method (the second column), MDAL method (the third column), ISD-SB method (the fourth column), JLISD-SB method (the fifth column), IISD-SB method (the sixth column). The PR value from first row to fourth row ranges from 30% to 60%, respectively. Noise level: $\sigma = 5$. Test image: cameraman.

However, in this paper, the looser stopping tolerance during the intermediate stages is not adopted in order for better showing the intermediate results. In such cases, (JL)ISD-SB might be several times longer than SB and in our experiments usually 2–3 times longer. Notice that for our problems, computational time is usually not a big concern, since it only usually takes minutes for our testing problems, and we are mainly focusing on the recovery quality and (JL)ISD-SB is performing much better than SB in this aspect.

IV. NUMERICAL EXPERIMENTS

In this section, we provide comparisons between the split bregman algorithm (Algorithm 1) for the analysis based approach (ASBA), and the iterative support detection split bregman (Algorithm 3) for the weighted analysis based approach (weighted ASBA) to show that the ISD, the multi-stage self-learning procedure can promote the performance of image inpainting. Moreover, we make comparisons with a wavelet frame based nonconvex algorithm for image inpainting, i.e., mean doubly augmented Lagrangian (MDAL) method in [7], where the authors proposed to penalize the l_0 norm of the wavelet frame coefficients instead, and they have demonstrated significant improvements of their algorithm over some commonly used l_1 minimization models in terms of

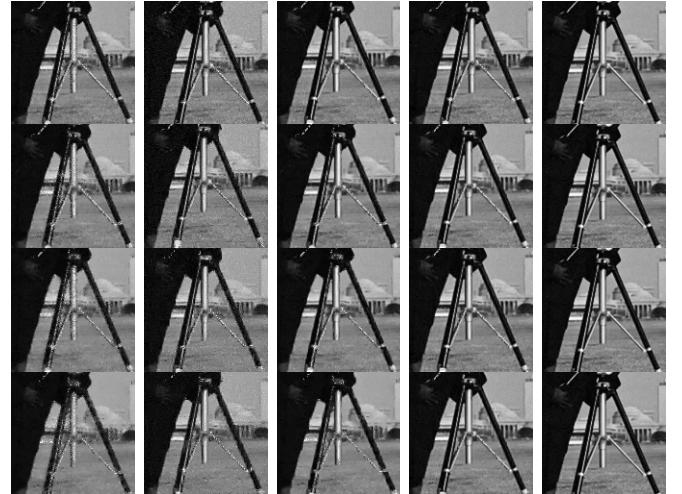


Fig. 10. The zoom in visual comparisons between SB method (the first column), MDAL method (the second column), ISD-SB method (the third column), JLISD-SB method (the fourth column), IISD-SB method (the fifth column). The PR value from first row to fourth row ranges from 30% to 60%, respectively. Noise level: $\sigma = 5$. Test image: cameraman.

image deblurring. We select 6 commonly used images for the tests. We compare the quality of the recoveries not only on the increasing of SNR and SSIM values (defined in the following

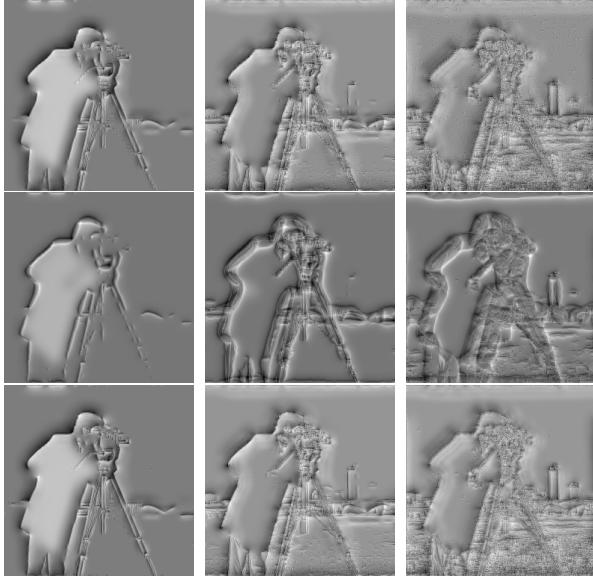


Fig. 11. The inverse wavelet frame transform to the support detection after each stage of ISD-SB (first row), JLISD-SB (second row), IISD-SB (third row), where the first, second, and the third columns are corresponding to the support maps after the first stage, second stage, and the third stage, respectively. PR = 30%. Test image: cameraman.



Fig. 12. The back-projection results after each stage of ISD-SB (first row), JLISD-SB (second row), IISD-SB (third row), which obtained by only reserving the large wavelet frame coefficients to the underlying true solution at the support detections, where the first, second, and the third columns are corresponding to the back-projection results after the first stage, second stage, and the third stage, respectively. PR = 30%. Test image: cameraman.

part), but also the visible detail improvements of recovered images. Corruptions by both pixels missing and ruleless lines are considered in our numerical experiments. We also test the stability of the involved algorithms by increasing the amount of the noise added to the observed images. All the experiments were performed under Windows 7 and MATLAB v7.10.0 (R2010a) running on a desktop with an Intel(R) pentium(R) CPU G640 (2.80GHz) and 2GB of memory.

A. Experiments Settings and Choices of Parameters

We select 6 test images as shown in Figure 3 which are “checkerboard”, “lena”, “cameraman”, “boat”, “bowl”, and “man”, respectively. In Figure 4, we show that both the magnitudes of sorted wavelet frame coefficients and sorted joint-level group magnitudes of wavelet frame coefficients of the 6 tested images have a property of fast decaying, which demonstrate that the fast decaying property is a mild assumption in practice.

The intensity of a pixel of these gray-scale images ranges from 0 to 255. We consider several different Gaussian noise levels with mean being 0 and the standard deviation σ ranging from 0 to 10. For the specific image inpainting problem, the operator A in (1) is a projection operator defined as

$$(Au)_i = \begin{cases} u_i, & i \in \Lambda \\ 0, & \text{otherwise} \end{cases}$$

where Λ is the domain where information of the underlying image u is known. We define the projection ratio (PR) as

$$\text{PR} = \frac{n - |\Lambda|}{n}$$

where $|\Lambda|$ is the cardinality of Λ , and n is equal to the number of total pixels. It is clear that the larger PR value indicates the more corrupted observed image, and the PR value ranges from 30% to 60% in our tests. In practice, since we do not know the original true solution, the weights are determined by the iterative support detection according to the intermediate recovered results, which can be understood as an adaptive learning process. In the following numerical experiments, we also give the ideal recovered results of our proposed model, i.e., we use the plain support detection strategy, but based on the underlying true solution. While we usually do not know the true solution in practice, here we just use it as a reference and named it as Ideal Iterative Support Detection Split Bregman (IISD-SB) method, as an ideal golden upper bound of the performance of ISD based methods.

The quality of the recovered image is measured by the SNR value defined as

$$\text{SNR} := 10\log_{10} \frac{\|u - \bar{u}^0\|_2}{\|u - u^0\|_2}$$

where u^0 is the original image and \bar{u}^0 is the mean intensity of u^0 . In addition, we also use another image quality assessment: Structural SIMilarity (SSIM) index between two images proposed in [62], which aims to be more consistent with human eye perception. Here we omit the detailed definitions for the limit of length of the paper.

Throughout the numerical experiments, the Split Bregman algorithm is terminated when one of the following situation is satisfied:

$$\min \left\{ \frac{\|u^k - u^{k-1}\|_2}{\|u^k\|_2}, \frac{\|Au^k - f\|_2}{\|f\|_2} \right\} < 5 \times 10^{-4}$$

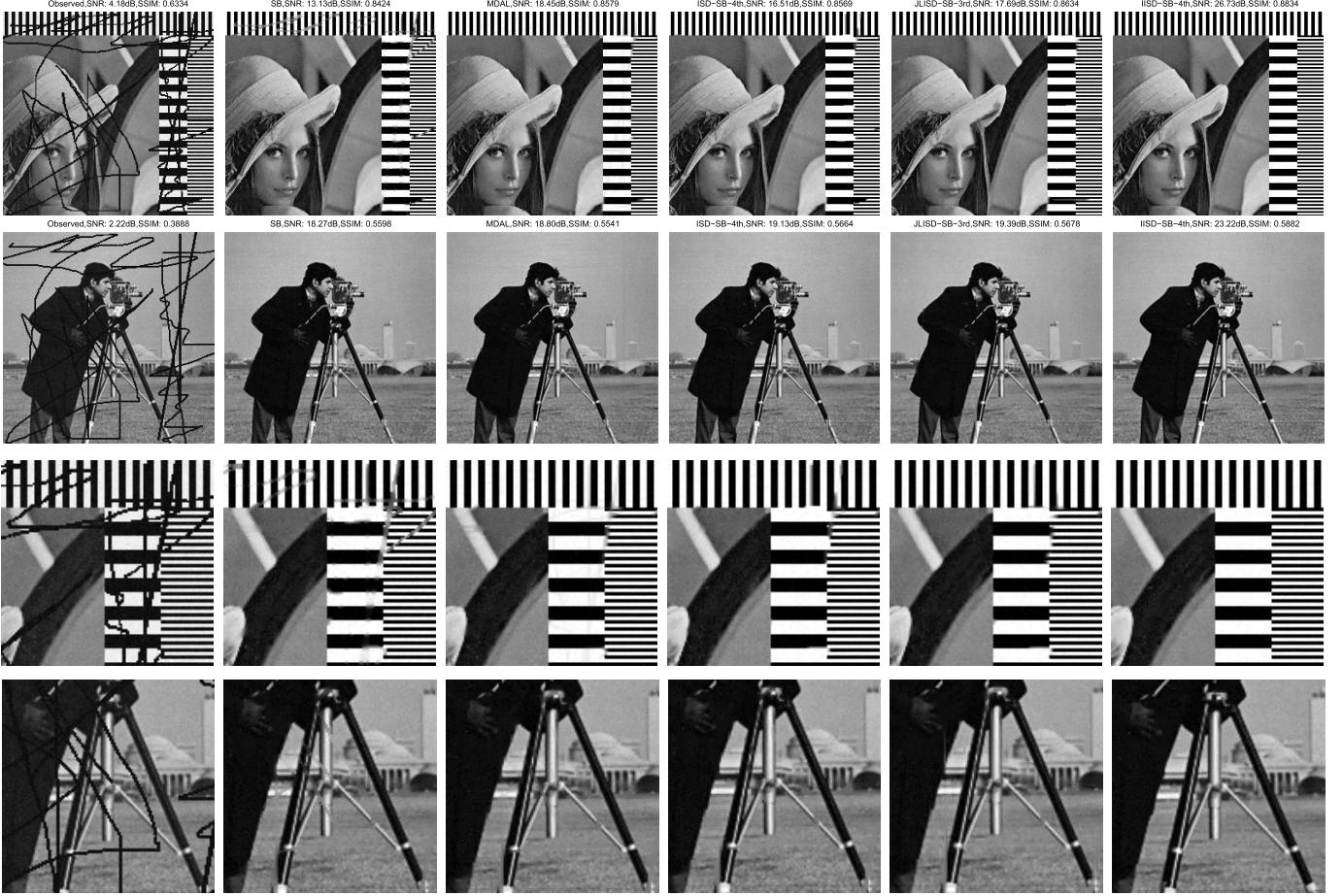


Fig. 13. The observed images are corrupted by ruleless lines (the first column). The recovered results of SB method (the second column), MDAL method (the third column), ISD-SB method (the fourth column), JLISD-SB method (the fifth column), IISD-SB method (the sixth column). The third and fourth rows show close-up views of the corresponding images in the first and second rows. Noise level: $\sigma = 5$. Test images: *lena* and *cameraman*.

or the iteration number reaches the prescribed maximal number, e.g., 50. We use the same inner iteration condition for ISD-SB method, JLISD-SB method, and IISD-SB method, i.e., the stopping criterion of 2(b) in Algorithm 3.

For all the cases tested, for SB method, we take the parameter $\lambda_{l,i,j}$ in (14) as follows: $\varrho_l = (\frac{1}{2})^l \beta$ with some carefully chosen scalar β .

According to our experience, the framelet decomposition level to be 4 is a commonly desirable choice for image inpainting, and thus we fix it to be 4 (i.e., $L = 4$). We select the optimal parameters μ and β for the different cases in terms of the best SNR and SSIM values. For our proposed ISD-SB method, JLISD-SB method and the ideal IISD-SB method, the parameters are set to be the same as SB method. Empirically, these parameters are not very sensitive to the types of images, pixel missing levels (i.e., PR values) and noise levels, which can be reflected by Figure 5 and 6, where we change the PR value and noise level, and observe the SNR and SSIM variation trends as the parameter μ and β vary. Here we just show the results of “*lena*” image of ISD-SB method for paper brevity, since the other test images have the similar conclusions. In addition, we set $\rho = 3$ in (26), $\eta = 11$ in (28) throughout the experiment tests. According to our experience, the number of stages is not often necessarily very large, like

around 4 in practice, and we also show the recovered results of the different choices of ρ and η .

For the wavelet frame based nonconvex MDAL algorithm, the default settings in [7] are used in the following comparisons (we refer the readers to [7] for more details), i.e., fix $\mu = 0.01$, $\gamma = 0.003$, and select the optimal parameter λ in terms of SNR and SSIM values. We believe our comparisons will be fair under these parameter settings.

B. Results and Discussions

In Tables I and II, we summarize the results of the five algorithms for all the 6 test images with noise level $\sigma = 5$ and $\sigma = 10$, respectively. We observe that the ISD-SB method, JLISD-SB method and IISD-SB method have an overall better performance than the SB method in terms of SNR and SSIM values. In addition, we observe that our proposed iterative support detection based split bregman methods can achieve much better performance over MDAL in terms of SNR and SSIM values in most cases. Notice that because MDAL works well for barcode-like patterns (binary bars) as mentioned in [7], MDAL sometimes achieves even better results for the “checkerboard” and “*lena*” images, which are of significant

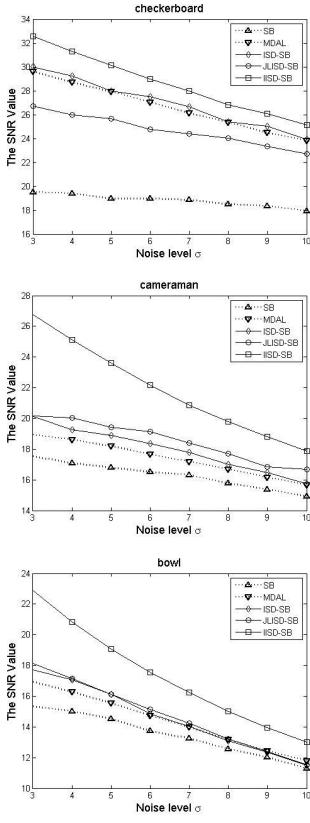


Fig. 14. Curves of SNR value v.s. noise level for the SB method, MDAL method, ISD-SB method, JLISD-SB method, and IISD-SB method. PR = 30%.

amount of barcode-like patterns. However, for other natural images, JLISD-SB mostly achieves significantly better results, partially due to the usage of the structure of the wavelet frame coefficients. In the following, we will give more plots to help demonstrate these points.

In Figures 7 and 8, we give the visual comparisons between SB, ISD-SB and JLISD-SB recovered results, as well as the intermediate results of each stage of ISD. The test image is ‘‘cameraman’’ here, the noise level is fixed as $\sigma = 5$, and the projection ratio (PR) ranges from 30% to 60%. Recall that the first stage of both ISD-SB and JLISD-SB method is to run the plain SB method, and we can observe that our proposed algorithm can bring gradual improvements as the ISD proceeds. Empirically, for most cases, ISD-SB and JLISD-SB method achieves the best final results when the maximum stage number reaches 4 and 3, respectively.

Now we show the recovered images of the five algorithms, i.e. SB, MDAL, ISD-SB, JLISD-SB and IISD-SB, respectively. Due to the page limit, we just show the cases listed in Tabel 1 in Figure 9, where the test image is ‘‘cameraman’’ here, since the other cases have the similar conclusions. In addition, for better visual comparisons, in Figure 10, we give the close-up views (zoom-ins) of results. We can observe that JLISD-SB brings significant visual enhancements in the sharp edges of the recovered images compared to SB method and MDAL method, as reflected by the SSIM values.

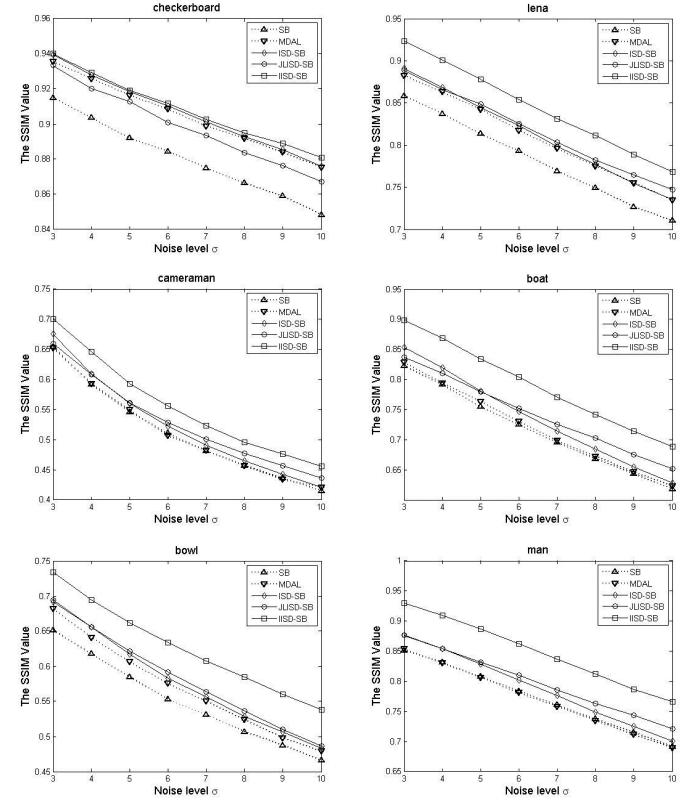


Fig. 15. Curves of SSIM value v.s. noise level for the SB method, MDAL method, ISD-SB method, JLISD-SB method, and IISD-SB method. PR = 30%.

To further study our proposed method, we also provide some intermediate detected support maps as ISD proceeds. We fix the PR as 30% and choose the test image ‘‘cameraman’’ here. In Figure 11, we show the pictures of inverse wavelet frame transform to the support detection (binary 0-1 coefficients, the coefficients on the support detections are 1, and the remainder are 0) after each stage of ISD-SB, JLISD-SB, and IISD-SB methods. In Figure 12, we provide the back-projection results after each stage of ISD-SB, JLISD-SB, and IISD-SB methods, which obtained by only reserving the large wavelet frame coefficients to the underlying true solution at the support detections. From the above observations, it is demonstrated again that the we indeed obtain more reliable true support detection as the ISD proceeds.

In Figure 13, we also present the inpainting results of the images corrupted by ruleless lines, where we only use the test images ‘‘lena’’ and ‘‘cameraman’’ as examples, since the other test images have the similar results. From the zoom-in visual comparisons, we can observe that ISD-SB, JLISD-SB, IISD-SB and MDAL methods achieve better inpainting performance compared to SB method in terms of better preserving the edges of these two test images. We note that for the image ‘‘lena’’, MDAL achieves a better SNR over ISD methods due to a significant portion of barcode-like patterns which are ideal for MDAL (see [7] for details). However, JLISD-SB achieves a better SSIM value.

In Figures 14 and 15, we also test the stability of the five algorithms. As the noise level increases, the

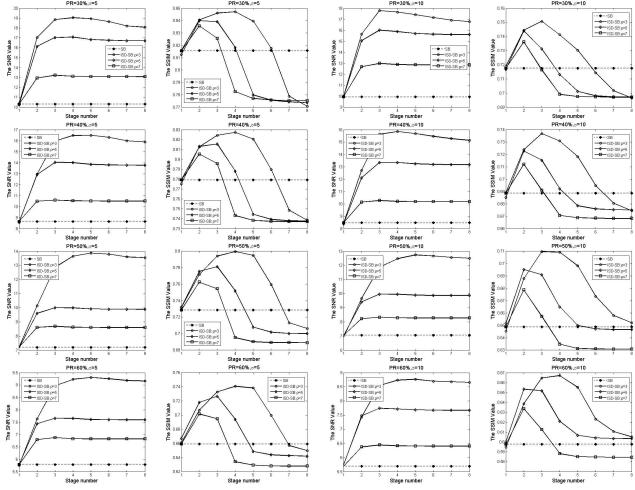


Fig. 16. Curves of SNR value and SSIM value v.s. ρ choices for the ISD-SB method. Test image: *lena*.

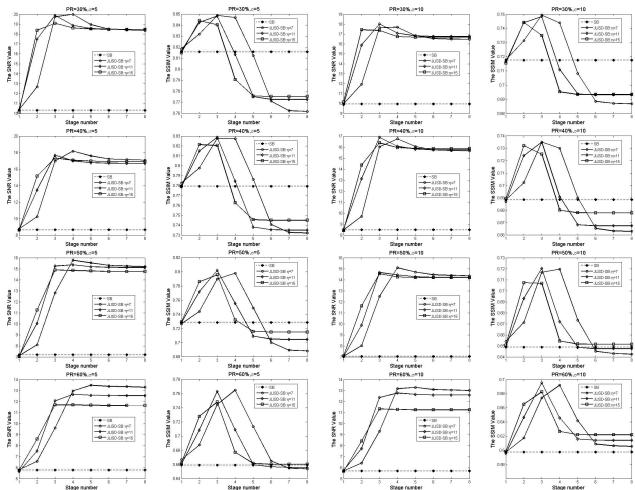


Fig. 17. Curves of SNR value and SSIM value v.s. η choices for the JLISD-SB method. Test image: *lena*.

degradation of recovered image quality is linear and ISD based methods have better performance in most cases, though the superiority of our proposed ISD-SB method, and JLISD-SB method over the SB method and MDAL method shrinks, because reliable support detection becomes more difficult to acquire as the noise level and/or the PR value increases.

In Figures 16 and 17, we present the different recovery results of ISD-SB method and JLISD-SB as the parameters ρ and η change, respectively. For the paper conciseness, we only choose the test image “*lena*” here, and fix the PR value as 30%, since the other cases have the similar conclusions. For *threshold*-ISD strategy (0-1 weighting scheme) belonging to greedy methods, the different options of ρ and η reflect the different greedy degree, showing the confident level that we detect the true supports. In particular, the more smaller the ρ and η we choose, the more greedy we act and the more confidence we have. We observe that we can achieve different best recovered results at different stages as the parameters ρ and η

change. Moreover, when the stage number s becomes too big, the results become worse instead, because the threshold value becomes very small, resulting into too many wrong detections, according to the formulas (26) and (28). In such cases, we can set up a lower bound of threshold according to certain prior information to avoid excessive support detections.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we propose the ISD-SB method and especially the JLISD-SB method for wavelet frame based image inpainting problems. The final result is obtained from a multistage process consisting of solving a series of weighted ASBA model, where *threshold*-ISD strategy is applied to the rough intermediate results to determine the adaptive binary weights.

ISD-SB method and especially JLISD-SB method can bring significant enhancements at the sharp edges of the recovered images compared to SB method and MDAL method, because the large wavelet frame coefficients reflect the singularities of the underlying true solution, and the *threshold*-ISD strategy sharpens the edges of the approximate solution by not thresholding the large coefficients and thus encouraging the edges to form in the recovery. More important, JLISD makes use of the joint sparsity property of the wavelet frame coefficients [63], [64] to further improve the recovery quality. Therefore future research includes studying specific image classes (including color images) and developing more effective corresponding support detection methods.

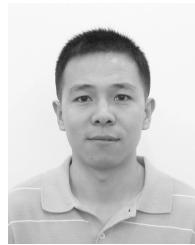
REFERENCES

- [1] R. H. Ghan, T. F. Chan, L. Shen, and Z. Shen, “Wavelet algorithms for high-resolution image reconstruction,” *SIAM J. Sci. Comput.*, vol. 24, no. 4, pp. 1408–1432, 2003.
- [2] J.-F. Cai, S. Osher, and Z. Shen, “Split Bregman methods and frame based image restoration,” *Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–367, 2009.
- [3] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, “Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA),” *Appl. Comput. Harmonic Anal.*, vol. 19, no. 3, pp. 340–358, 2005.
- [4] J.-L. Starck, M. Elad, and D. L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach,” *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
- [5] Z. Shen, “Wavelet frames and image restorations,” in *Proc. Int. Congr. Math.*, vol. 4, 2010, pp. 2834–2863.
- [6] B. Dong and Z. Shen, “MRA-based wavelet frames and applications,” in *The Mathematics of Image Processing* (IAS Lecture Notes Series). Salt Lake City, UT, USA: Park City Mathematics Institute, 2010.
- [7] B. Dong and Y. Zhang, “An efficient algorithm for ℓ_0 minimization in wavelet frame based image restoration,” *J. Sci. Comput.*, vol. 54, nos. 2–3, pp. 350–368, 2013.
- [8] Y. Zhang, B. Dong, and Z. Lu, “ ℓ_0 minimization for wavelet frame based image restoration,” *Math. Comput.*, vol. 82, no. 282, pp. 995–1015, 2013.
- [9] I. F. Gorodnitsky and B. D. Rao, “Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm,” *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.
- [10] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, “ k -t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI,” *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, 2009.
- [11] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [12] A. Chambolle and P.-L. Lions, “Image recovery via total variation minimization and related problems,” *Numer. Math.*, vol. 76, no. 2, pp. 167–188, 1997.

- [13] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 462–491, 2010.
- [14] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, Sep. 2010.
- [15] W. Lu and N. Vaswani, "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 182–196, Jan. 2012.
- [16] L. Jacques, "A short note on compressed sensing with partially known signal support," *Signal Process.*, vol. 90, no. 12, pp. 3308–3312, 2010.
- [17] M. P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1122–1134, Feb. 2012.
- [18] R. E. Carrillo, L. F. Polania, and K. E. Barner, "Iterative algorithms for compressed sensing with partially known support," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2010, pp. 3654–3657.
- [19] R. Carrillo, L. Polania, and K. Barner, "Iterative hard thresholding for compressed sensing with partially known support," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2011, pp. 4028–4031.
- [20] W. Guo and W. Yin, "Edge guided reconstruction for compressive imaging," *SIAM J. Imag. Sci.*, vol. 5, no. 3, pp. 809–834, 2012.
- [21] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [22] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1233, 2006.
- [23] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*. Providence, RI, USA: AMS, 2001.
- [24] G. Sapiro, *Geometric Partial Differential Equations and Image Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2001.
- [25] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*. New York, NY, USA: Springer-Verlag, 2003.
- [26] T. Chan, S. Esedoglu, F. Park, and A. Yip, "Total variation image restoration: Overview and recent developments," in *Handbook of Mathematical Models in Computer Vision*. New York, NY, USA: Springer-Verlag, 2006, pp. 17–31.
- [27] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. New York, NY, USA: Springer-Verlag, 2006.
- [28] R. Chan, L. Shen, and Z. Shen, "A framelet-based approach for image inpainting," *Res. Rep.*, vol. 4, p. 325, 2005.
- [29] J.-F. Cai, S. Osher, and Z. Shen, "Linearized Bregman iterations for frame-based image deblurring," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 226–252, 2009.
- [30] J.-F. Cai, B. Dong, S. Osher, and Z. Shen, "Image restorations: Total variation, wavelet frames and beyond," *J. Amer. Math. Soc.*, vol. 25, no. 4, pp. 1033–1089, 2012.
- [31] A. Ron and Z. Shen, "Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator," *J. Funct. Anal.*, vol. 148, no. 2, pp. 408–447, 1997.
- [32] I. Daubechies, *Ten Lectures on Wavelets* (CBMS-NSF Regional Conference Series in Applied Mathematics). Philadelphia, PA, USA: SIAM, 1992.
- [33] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames," *Appl. Comput. Harmonic Anal.*, vol. 14, no. 1, pp. 1–46, Jan. 2003.
- [34] I. Daubechies, G. Teschke, and L. Vese, "Iteratively solving linear inverse problems under general convex constraints," *Inverse Problems Imag.*, vol. 1, no. 1, p. 29, 2007.
- [35] M. J. Fadili and J.-L. Starck, "Sparse representations and Bayesian image inpainting," in *Proc. Int. Conf. SPARS*, vol. 5, 2005.
- [36] M. J. Fadili, J.-L. Starck, and F. Murtagh, "Inpainting and zooming using sparse representations," *Comput. J.*, vol. 52, no. 1, pp. 64–79, 2009.
- [37] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, Aug. 2003.
- [38] M. A. T. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 2, Sep. 2005, pp. II-782–II-785.
- [39] J. Cai, R. Chan, L. Shen, and Z. Shen, "Convergence analysis of tight framelet approach for missing data recovery," *Adv. Comput. Math.*, vol. 31, nos. 1–3, pp. 87–113, Oct. 2009.
- [40] J. Cai, R. Chan, and Z. Shen, "Simutaneous cartoon and texture inpainting," *Inverse Problems Imag.*, vol. 4, no. 3, pp. 379–395, 2010.
- [41] T. Goldstein and S. Osher, "The split Bregman algorithm for L1 regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
- [42] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.
- [43] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput. Math. Appl.*, vol. 2, no. 1, pp. 17–40, 1976.
- [44] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.
- [45] J. Eckstein and D. P. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, nos. 1–3, pp. 293–318, 1992.
- [46] M. R. Hestenes, "Multiplier and gradient methods," *J. Optim. Theory Appl.*, vol. 4, no. 5, pp. 303–320, 1969.
- [47] M. J. D. Powell, "A method for nonlinear constraints in minimization problems," in *Optimization*, R. Fletcher, Ed. New York, NY, USA: Academic, 1969, pp. 283–298.
- [48] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. Philadelphia, PA, USA: SIAM, 1989.
- [49] E. Esser, "Applications of Lagrangian-based alternating direction methods and connections to split Bregman," Dept. Math., UCLA, Los Angeles, CA, USA, CAM Rep., 2009.
- [50] X.-C. Tai and C. Wu, "Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model," *Scale Space and Variational Methods in Computer Vision*. Berlin, Germany: Springer-Verlag, 2009, pp. 502–513.
- [51] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [52] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, 2006.
- [53] P. Mrázek and J. Weickert, "Rotationally invariant wavelet shrinkage," in *Pattern Recognition*. Berlin, Germany: Springer-Verlag, 2003, pp. 156–163.
- [54] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk, "On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDES," *SIAM J. Numer. Anal.*, vol. 42, no. 2, pp. 686–713, 2005.
- [55] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imag. Sci.*, vol. 1, no. 3, pp. 248–272, 2008.
- [56] S. Becker, J. Bobin, and E. J. Candès, "NESTA: A fast and accurate first-order method for sparse recovery," *SIAM J. Imag. Sci.*, vol. 4, no. 1, pp. 1–39, 2011.
- [57] X. Zhang, M. Burger, and S. Osher, "A unified primal-dual algorithm framework based on Bregman iteration," *J. Sci. Comput.*, vol. 46, no. 1, pp. 20–46, 2011.
- [58] T. Zhang, "Analysis of multi-stage convex relaxation for sparse regularization," *J. Mach. Learn. Res.*, vol. 11, pp. 1081–1107, Mar. 2010.
- [59] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
- [60] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.
- [61] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. 33rd Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar./Apr. 2008, pp. 3869–3872.
- [62] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [63] C. Chen and J. Huang, "Compressive sensing MRI with wavelet tree sparsity," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran & Associates Inc., 2012, pp. 1124–1132.
- [64] C. Chen and J. Huang, "The benefit of tree sparsity in accelerated MRI," *Med. Image Anal.*, to be published.
- [65] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.



Liangtian He (M'–) is currently pursuing the Ph.D. degree with School of Mathematical Science, University of Electronic Science and Technology of China, Chengdu, China. His research interests are in image and signal processing (in particular, restoration, compression, and compressive sensing), computer vision, pattern recognition, and machine learning.



Yilun Wang received the B.S. degree from the Dalian University of Technology, Dalian, China, in 2002, the M.S. degree from Fudan University, Shanghai, China, in 2005, and the Ph.D. degree from Rice University, Houston, TX, USA, in 2009, respectively, all in computational and applied mathematics. From 2009 to 2012, he was a Post-Doctoral Researcher with Cornell University, Ithaca, NY, USA. Since 2012, he has been an Associate Professor with the School of Mathematical Sciences, University of Electronic Science and Technology of

China, Chengdu, China. He is working on developing fast algorithms to find the optimal sparse representation, with applications in visual computing, machine learning, and data mining to deal with high-dimensionality. He has authored several papers in high-ranked journals, such as the *SIAM Journal on Imaging Sciences*, *Information Sciences*, and the *Journal of Agricultural, Biological, and Environmental Statistics*. One of his papers has been cited over 640 times according to the Google Scholar.