Journal of Scientific Computing manuscript No.
(will be inserted by the editor)

Data-driven Multi-scale Non-local Wavelet Frame Construction and Image Recovery

Yuhui Quan · Hui Ji · Zuowei Shen

Received: date / Accepted: date

Abstract By assuming that images of interest can be sparsely modelled by some transform, the sparsity-based regularization has been one promising approach for solving many ill-posed inverse problems in image recovery. One often-used type of systems for sparsifying images is wavelet tight frames, which can efficiently exploit the sparse nature of local intensity variations of images. However, existing wavelet frame systems lack the capability of exploiting another important image prior, i.e., the self-recursion of local image structures in both spatial and scale domain. Such a self-recursion prior of image structures has led to many powerful non-local image restoration schemes with impressive performance. This paper aims at developing a scheme for constructing a non-local wavelet frame or wavelet tight frame that is adaptive to the input image. The proposed multi-scale non-local wavelet frame allows the resulting regularization simultaneously exploits both the sparse prior of local variations of image intensity and the global self-recursion prior of image structures in spatial domain and across scales. Based on the proposed construction scheme, a powerful regularization method is developed for solving image deconvolution problem. The experiments showed that the results from the proposed regularization method are compared favorably against that from several popular image restoration methods.

Keywords non-local scheme · ℓ_1-norm minimization · image restoration · wavelet frames

Mathematics Subject Classification (2010) 68U10 and 65F22 and 65T60

1 Introduction

Most image restoration tasks are about recovering a high-quality image from its partial and noisy measurement. The measurement is often modeled by a linear

Y. Quan and H. Ji and Z. Shen
Department of Mathematics,
National University of Singapore
Singapore 119076
E-mail: {matquan,matjh,matzuows}@math.nus.edu.sg

operator $A \in \mathbb{R}^{M\times N}$ applied to the input image:

$$g = Af + n, \tag{1}$$

where $f \in \mathbb{R}^N$ denotes the image for recovery, $g \in \mathbb{R}^M$ denotes the available measurement of f and $n \in \mathbb{R}^M$ denotes the measurement noise. For image noise removal, the matrix A is an identity matrix. For image de-blurring, the matrix A represents a convolution process by a low-pass filter that blurs images. Most of these image recovery problems are ill-posed inverse linear problems in which straightforward solutions in term of matrix inversion lead to useless solutions dominated by noise. To solve these ill-posed image restoration problems, certain prior information of the image of interest is needed to regularize the recovery process. In recent years, the sparsity-based prior has been playing a very important role in the recent development of effective image restoration algorithms. The sparsity-based image prior assumes that the image of interest is compressible in some transform domain, that is, most of the important information of the image can be kept by using few transform coefficients. Thus, the image recovery process can be regularized by minimizing the functional that prompts the sparsity of the solution in the transform domain. One such convex sparsity promoting functional is the ℓ_1 norm of transform coefficients of the solution. Clearly, the effectiveness of these sparsity-based regularization methods is highly dependent on how efficiently the chosen transform can sparsify images of interest.

In the last few decades, the orthonormal *wavelet bases* [1,2] have been widely used in many image processing tasks. For image recovery, redundant over-complete systems are often more preferred, as it is empirically observed that the results from over-complete systems tend to be more visually appealing with less artifacts. In recent years, as a generalization of orthonormal wavelet bases, *wavelet tight frames* [3,4,5] have been used to sparsify images of interest in many applications. Many types of wavelet tight frames have been proposed and are applied in various image restoration tasks. For example, shift-invariant wavelet system for image de-noising [6], curvelet [7] and its applications in image de-noising and de-blurring [8,9], bandlet [10] for image approximation and compression [11], framelet [3,4] and its applications in many image restoration tasks including in-painting, deblurring and etc (see e.g. [12,5,13,14]). It is noted that the widely used total variation [15] based variational methods are deeply connected with the ℓ_1 norm based regularization under spline framelets. By choosing parameters properly, the analysis approach using spline frames [5] indeed can be viewed as a sophisticated discretization of variational methods involving the total variation penalties (see [16] for more details). The construction of these tight frames relies on certain functional assumptions of images which sometimes are invalid for the particular types of images, an alternative approach is then to adaptively choose the redundant system optimized for the input image. Several adaptive schemes have been developed in recent years for sparsely modelling images. For example, the K-SVD method [17] learns an over-complete system (frame) from the input image, and the data-driven tight frame construction method [18] constructs a shift-invariant tight frame adaptive to the input image.

The elements of the wavelet frames discussed above are all locally supported in spatial domain. Then, for each image pixel, the related wavelet frame coefficients measure local image gradients of different degrees in its local neighborhood. In other words, the sparsity prior of images under an existing wavelet frame system

only refers to the sparse nature of local variations of image intensity. Thus, we may call the existing wavelet frame based regularization and total variation (TV) based regularization a *local* approach. The local approach works well on cartoon-type regions of the image, but leaves plenty of room for improvement on textural regions. An alternative approach is the non-local approach which relates many image pixels that may be far away in spatial domain. The non-local approach is built on one observation often seen in natural images: image structures of small image regions tend to repeat themselves in spatial domain. Such a prior is non-local in the sense that image pixels from far way may be related to each other. There are two types of non-local schemes proposed to use such a non-local prior. One is the so-called *non-local mean* first introduced in [19] for image de-noising and extended to solve other inverse problems in image processing; see e.g. [20,21,22,23]. The non-local mean approach introduces a non-local operator, e.g. a weighted averaging filter, in the variational model to explore the spatial redundancy. Another is the patch-based approach which groups the similar patches together followed by a collaboratively filtering; see e.g. [24,25,26,27]. One well-known approach is the BM3D method for image de-noising [24] and image deconvolution [25].

The local approaches and the non-local approaches discussed above both have their own advantages and disadvantages. For solving linear inverse problems in image recovery, most local approaches can be expressed by a variational formulation. Their performances are also stable and consistent over many types of images, particularly for images dominated by cartoon-type image structures such as medical images. However, these local approaches do not work very effectively for natural images with complex texture regions, as the local variations of these regions are not sparse. The non-local mean approach addressed such a weakness by introducing a non-local operator in the variational formulation, which leads to better performance than local approaches. However, the performance of the non-local mean approaches is still not comparable to that of some patch-based non-local approaches, e.g. the BM3D method. The performance of some patch-based non-local methods are very impressive in image de-noising and image deconvolution. By applying a 3D filter on the stack of matched image patches, the BM3D method considered both the local sparsity prior of image intensity and the global self-recursion prior of image structures. However, it did not introduce the mechanism to deal with the images which lack self-recursive image structures, such as fingerprint images or medical images. Thus, by wrongly imposing global similarity prior on image patches, the BM3D method is likely to yield poor results for these type of images.

Aiming at having the best of both local and non-local approaches, this paper proposed a scheme of constructing non-local wavelet frames (or tight frames) adaptive to the input image, which will allow the resulting sparsity-based regularization simultaneously to exploit three image priors: (i) sparsity prior of local intensity variations; (ii) self-recurrence prior of local image structures in spatial domain; and (iii) self-recurrence prior of local image structures across scales. The main differences between our proposed approach and the existing approaches are summarized as follows.

(i) In contrast to existing wavelet frames, the proposed non-local wavelet frame is composed of the concatenation of a multi-scale frame (or tight frame) with non-local elements and a multi-scale tight frame with localized elements. Such

a frame can facilitate the effective usage of all three image priors mentioned above. Built upon the proposed construction scheme of non-local wavelet frames, a variational image recovery model is developed to exploit the sparsity prior of local variations and the self-recurrence of local image structures in spatial domain and across multiple scales.

(ii) Different from the non-local mean methods, the proposed method utilized the global similarity prior in a way that is essentially the same as patch-based methods. As a result, the better performance than the non-local mean methods is observed in experiments.

(iii) Different from the existing patch-based methods (e.g. [24,25,26,27]), the proposed variational formulation works for general image restoration problems, and is adaptive to the images with or without strong self-recursion prior. In addition, the non-local component in the proposed frame is built on multi-scale wavelet tight frames. The inherent multi-scale property of wavelet tight frames allows the usage of the self-recurrence prior of local image structures across scales, which leads to further performance improvement for image recovery.

The rest of the paper is organized as follows. In Section 2, we give a brief review on frames, wavelet frames, tight frames, local and non-local image restoration schemes . The main results are presented in Section 3. Section 4 is devoted to the experimental evaluation of the proposed image recovery method.

2 Preliminaries and previous work

2.1 Wavelet frames, tight frames and image recovery

We first present here some basics of frames and tight frames in a Hilbert space H. Interested readers are referred to [4,5,28] for more details. Let $\mathbb{Z}$ denote the set of all integers, $\mathbb{Z}_M$ denote the integer set $\{1, 2 \ldots, M\}$, and let $\mathbb{Z}^2$ denote the set of all 2D integers. Let $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the usual inner product and norm of a Hilbert space $\mathcal{H}$. A sequence $\{\phi_n\}_{n\in\mathbb{Z}} \subset \mathcal{H}$ is a frame for $\mathcal{H}$ if there exist two positive constants a and b such that

$$a\|f\|_2^2 \leq \sum_{n\in\mathbb{Z}} |\langle \phi_n, f\rangle|^2 \leq b\|f\|_2^2, \quad \forall\, f \in \mathcal{H}.$$

A frame $\{\phi_n\}_{n\in\mathbb{Z}}$ is called a tight frame for H when $a = b = 1$. There are two operators associated with a given frame $\{\phi_n\}_{n\in\mathbb{Z}}$: the analysis operator W defined by

$$W : f \in \mathcal{H} \longrightarrow \{\langle f, \phi_n\rangle\} \in \ell^2(\mathbb{Z})$$

and its adjoint operator W^*, also called the synthesis operator, defined by

$$W^* : \{a_n\} \in \ell^2(\mathbb{Z}) \longrightarrow \sum_n a_n \phi_n \in \mathcal{H}.$$

The concatenation of these two operators forms a so-called frame operator $S = W^*W$ given by

$$S : f \in \mathcal{H} \longrightarrow \sum_{n\in\mathbb{Z}} \langle f, \phi_n\rangle \phi_n.$$

Thus, a sequence $\{\phi_n\} \subset \mathcal{H}$ forms a frame if and only if $aI \leq S \leq bI$ and it forms a tight frame if and only if $S = I$, where $I : \mathcal{H} \longrightarrow \mathcal{H}$ is the identical operator. Frame can be viewed as a generalization of Riesz Basis and tight frame is a redundant system that generalizes orthonormal basis. Indeed, a given tight frame $\{\phi_n\}_{n\in\mathbb{Z}}$ has the same perfect reconstruction property as orthonormal basis:

$$f = \sum_{n\in\mathbb{Z}} \langle f, \phi_n\rangle \phi_n, \quad \forall\, f \in \mathcal{H}, \tag{2}$$

and it becomes an orthonormal basis if $\|\phi_n\| = 1$ for all ϕ_n.

Wavelet tight frames are arguably the most often used frames in signal/image processing. A wavelet tight frame for $L_2(\mathbb{R})$ is a system formed by the shifts and dilations of a finite set of generators $\Psi = \{\psi_1, \ldots, \psi_m\} \subset L_2(\mathbb{R})$:

$$X(\Psi) = \{2^{j/2}\psi_\ell(2^j \cdot -k), \quad 1 \leq \ell \leq m, j \in \mathbb{Z}, k \in \mathbb{Z}\}.$$

Framelets are the wavelet tight frames constructed via multi-resolution analysis. The construction of framelets starts with a scaling function ϕ with $\widehat{\phi}(0) = 1$ that satisfies the following refinable equation $\widehat{\phi}(2\cdot) = \widehat{a_0}\widehat{\phi}$, where $\widehat{\phi}$ is the Fourier transform of ϕ, and $\widehat{a_0}$ is a 2π-periodic trigonometric polynomial

$$\widehat{a_0}(\omega) := \sum_{k\in\mathbb{Z}} a_0(k) e^{-ik\omega}$$

with $\widehat{a_0}(0) = 1$. Then the generators $\{\psi_1, \ldots, \psi_m\}$ is defined by $\widehat{\psi_\ell} = \widehat{a_\ell}\widehat{\phi}$, $\quad 1 \leq \ell \leq m$. The so-called Unitary Extension Principle ([3]) states that $X(\Psi)$ forms a tight frame provided that $\phi \in L^2(\mathbb{R})$ and

$$\sum_{\ell=0}^{m} \widehat{a}_\ell(\omega)\overline{\widehat{a_\ell}(\omega + \pi\gamma)} = \delta_\gamma, \quad \gamma = 0, 1. \tag{3}$$

For instance, the linear B-spline framelet often used in image recovery (see e.g. [12, 29]) has two generators and the associated masks $\{a_0, a_1, a_2\}$ are

$$a_0 = \frac{1}{4}[1, 2, 1]; \quad a_1 = \frac{\sqrt{2}}{4}[1, 0, -1]; \quad a_2 = \frac{1}{4}[-1, 2, -1]. \tag{4}$$

The framelet system for $L_2(\mathbb{R}^2)$ can be obtained by taking the tensor product of univariate framelets. In the discrete setting, for a given frame $\{\phi_n\}_{n=1}^M \subset \mathbb{R}^N$, the matrix representation of the associated analysis operator is $W = (\phi_1, \phi_2, \ldots, \phi_M)^\top$ and the synthesis operator is its transpose $W^\top$. Clearly, $\{\phi_n\}_{n=1}^M$ forms a frame for $\mathbb{R}^N$ if and only if $W^\top W$ is non-singular and it forms a tight frame if and only if $W^\top W = I_N$, where I_N is the N-by-N identical matrix.

The wavelet tight frame for $\mathbb{R}^N$ can be constructed from the masks associated with the framelets for the space of continuum. For simplicity, only L-level un-decimal wavelet tight frame system for $\mathbb{R}^N$ is introduced. Let a_0 denote the mask associated with the scaling function and $\{a_1, a_2, \ldots, a_m\}$ denote the masks associated with other framelets. Then, the low-pass filter $a_0^{(\ell)}$ at level ℓ is define as

$$a_0^{(\ell)} = a_0^{\downarrow_1} * a_0^{\downarrow_2} * \cdots * a_0^{\downarrow_{2^{\ell-1}}}. \tag{5}$$

where $*$ denotes the discrete convolution operator and

$$a_k^{\downarrow_{2^\ell}} == [\ldots, a_k(-2), \overbrace{0, \ldots, 0}^{2^\ell-1}, a_k(-1), \overbrace{0, \ldots, 0}^{2^\ell-1}, a_k(0), \overbrace{0, \ldots, 0}^{2^\ell-1}, a_k(1), \overbrace{0, \ldots, 0}^{2^\ell-1}, a_k(2), \ldots],$$

for $k \in \mathbb{Z}$. Then, for the level $\ell = 1, \ldots, L-1$, the corresponding filter set is $\{a_k^{(\ell)}\}_{k=1}^m$ where

$$a_k^{(\ell)} = a_0^{(\ell-1)} * a_k^{\downarrow_{2^\ell}}. \tag{6}$$

For the L-th level, the filters are $\{a_k^{(L)}\}_{k=1}^m \cup a_0^{(L)}$.

For a given filter a of finite length, let the N-by-N matrix, denoted by $\mathcal{S}_a$, be the Toeplitz-plus-Hankel matrix that represents the convolution operator by the mask a under Neumann boundary condition (see [30]). Then, the analysis operator of a L-level discrete wavelet tight frame can be expressed as

$$W = [W_1^\top, W_2^\top, ..., W_L^\top]^\top, \tag{7}$$

where W_ℓ denotes the ℓ-th level analysis operator defined by

$$W_\ell = [\mathcal{S}_{a_1^{(\ell)}}^\top, \mathcal{S}_{a_2^{(\ell)}}^\top, \ldots, \mathcal{S}_{a_m^{(\ell)}}^\top]^\top, \quad \ell = 1, 2, \ldots, L-1, \tag{8}$$

and $W_L = [\mathcal{S}_{a_1^{(L)}}^\top, \mathcal{S}_{a_2^{(L)}}^\top, \ldots, \mathcal{S}_{a_m^{(L)}}^\top, \mathcal{S}_{a_0^{(L)}}^\top]^\top$. The rows of W forms a multi-level tight frame for $\mathbb{R}^N$ and its transpose $W^\top$ is the synthesis operator. The perfect reconstruction property of tight frame in matrix representation can be expressed as $W^\top W = I_N$.

Image restoration is about solving the ill-posed linear system (1) with some additional regularization on the solution. There are several approaches for wavelet frame based regularization, namely the synthesis based approach, the analysis based approach and the balanced approach; see [31,5] for more details. Here we focus on the analysis based approach which solves (1) via solving the following minimization model:

$$\underset{f}{\operatorname{argmin}} \frac{1}{2}||g - Af||_2^2 + ||\operatorname{diag}(\lambda)Wf||_1, \tag{9}$$

where W is the analysis operator of some discrete wavelet frame system, and λ is the weighting vector. In order to successfully recover the image of interest from (1) via solving (9), the image of interest should have a good sparse approximation under the discrete wavelet frame system. Many efficient algorithms have been developed in recent years to solve (9), e.g. the split Bregman iteration method [32, 33].

2.2 Related non-local image restoration schemes

There has been an abundant research literature on the non-local methods for various image restoration tasks. In this section, we will only discuss the most related work. The basic idea of most non-local methods in image processing is to recover the degraded pixel using other similar pixels. The non-local means (NLM) method is first introduced by Buades *et al.* [19] for image de-noising. In

the continuous setting, the non-local means method introduces a neighborhood de-noising filter applied to the image:

$$NL_f(x) := (\int_\Omega \omega_f(x,y)dy)^{-1} \int_\Omega \omega_f(x,y)f(y)dy. \tag{10}$$

where f is the reference image and ω_f is the weight function given by

$$\omega_f(x,y) = \exp(-\frac{G_\sigma * |f(x+\cdot) - f(y+\cdot)|^2(0)}{h^2}),$$

where G_σ is the Gaussian kernel with s.t.d. σ and h is the filtering parameter associated with the noise level. It is seen that the weights are significant only when the patch around y is similar to the corresponding patch around x. As a result, the self-similarity of image patches are used for noise reduction. The non-local means method is extended in [22] to address more general image restoration problems such as image deconvolution:

$$\text{argmin}_f \frac{\lambda}{2} \int_\Omega (k * f - g)^2 + \int_\Omega \int_\Omega \sqrt{(f(x) - f(y))^2 \omega_f(x,y)} dydx. \tag{11}$$

The regularization term in the above minimization clearly is a generalization of the TV based regularization by applying a neighborhood de-noising filter on TV measures.

A representative patch-based non-local scheme is the BM3D method [24,25]. The basic steps of the BM3D de-noising method [24] are as follows. First, similar image blocks are grouped and stacked in a 3D array based on the sum-of-squares distance function between different image patches. Then a collaborative filtering is carried out on each 3D image stack to suppress noise. For example, a shrinkage in 3D transform domain such as wavelet shrinkage or Wiener filtering. The de-noised image is then synthesized from the de-noised patches after inverting 3D transform. The result is then further refined by iteratively doing the patch grouping and collaboratively filtering. For image deconvolution, the BM3D deblurring method [25] takes a two-step approach. (i) The first step is to estimate an initial de-blurred result using the regularized inverse in discrete Fourier domain, followed by a de-noising process using the BM3D method with collaborative hard thresholding. Then this initial de-blurred result is used to synthesize a blurred image with better signal-to-noise ratio. (ii) Taking this re-estimated blurred image as the input, the second step is to re-estimate the result using the regularized Wiener filter, followed by a de-noising process using the BM3D method with collaborative Wiener filtering. In [26], a variational formulation of the the BM3D method is derived for general image recovery. The BM3D method is a very effective image restoration scheme with state-of-the-art performance.

3 Main results

Image restoration is about estimating the true image f by solving the linear system (1). In most image restoration problems, the matrix A in (1) is an ill-conditioned or non-invertible matrix. By assuming that the frame coefficient vector of the true solution f tends to be sparse, one popular sparsity-based approach is the

so-called analysis based approach (e.g. [31,5,32,33]), which estimates f by solving the following ℓ_1 norm related minimization model:

$$\underset{f}{\operatorname{argmin}} \Phi(g - Af) + ||\operatorname{diag}(\lambda) Df||_1, \tag{12}$$

where λ is some regularization parameter vector and D is the analysis operator of some frame system. There are two terms in the above minimization. The first term is the fidelity term and the functional Φ is set as $\frac{1}{2}\|\cdot\|_2^2$ when assuming only i.i.d. Gaussian white noise. The second term is the weighted ℓ_1 norm of coefficients which approximately measures the sparsity of the canonical frame coefficient vector Df. The performance of the model (12) largely depends on how well the analysis operator D can sparsify the true image f.

In this section, we first present a scheme for constructing the analysis operator D of non-local multi-scale wavelet frames and tight frames for $\mathbb{R}^N$, under which the regularization method such as (12) can exploit both the local sparsity prior of images in wavelet tight frame domain and the self-recursion prior of image structures in global spatial domain and across multiple scales. Then, built on the introduced non-local multi-scale wavelet frame system, a numerical method is developed for solving general image restoration problems.

3.1 Construction of multi-scale non-local wavelet frames

Many wavelet tight frames have been used for image restoration, including translation-invariant wavelets system [6] and spline framelets [12,13,14]. In the discrete case, these wavelet tight frame systems are generated by the shifts of several filters of small localized support, which are very suitable for sparsely approximating cartoon-type image regions composed of sparsely distributed edges. However, their effectiveness significantly decreases when approximating textural regions composed of dense small edge segments. For these textural regions, a more efficient approach is to represent these regions by themselves either in the same scale or in different scales, since these image structures are likely to repeat themselves in spatial domain or across different scales. Notice that if we represent these image regions under a wavelet tight frame, the associated wavelet tight frame coefficients will also have the same self-recursion property in spatial domain and across different scales. Thus, considering likely self-recursions of wavelet tight frame coefficients in spatial domain and across different scales, we propose a (tight) frame system whose analysis operator is of the form

$$D(f) = \frac{1}{\sqrt{2}} \begin{pmatrix} I \\ J(f) \end{pmatrix} W = \frac{1}{\sqrt{2}} \begin{pmatrix} W \\ J(f)W \end{pmatrix} \tag{13}$$

where $W \in \mathbb{R}^{M\times N}$ is the analysis operator associated with some existing multi-level wavelet frame system (e.g. 2-level shift-invariant cubic spline framelet [3]), $I \in \mathbb{R}^{M\times M}$ is the identity matrix, and $J(f) \in \mathbb{R}^{M\times M}$ is the operator that encodes the self-recursion property of wavelet tight frame coefficients Wf in spatial domain and in multiple scales.

The proposed frame system of the form (13) is composed of two systems: one is some existing discrete wavelet tight frame system denoted by W, and the other

is the non-local version of W generated by the multiplication of a linear operator $J(f)$. Since the discrete wavelet tight frame systems are generated by the shifts of the finitely-supported wavelet filters, the non-zero elements of each row of the matrix W are then agglomerated into some local interval. Thus, we may call the tight frame formed by the rows of W a *local* wavelet tight frame. The operator $J(f)$ is used for relating the wavelet frame coefficients corresponding to the same image structure, which often are not agglomerating in neighboring regions but rather are spreading out over the whole image. Thus the operator $J(f)$ can be viewed as a non-local operator to relate wavelet frame coefficients that may be spatially far away from each other, and the rows of $J(f)W$ can be viewed as a non-local version of discrete wavelet frame system generated by W. Therefore, we may call the frame (tight frame) of the form (13) a *non-local* wavelet frame (tight frame).

Proposition 1 *Let $W \in \mathbb{R}^{M\times N}$ be a tight frame system satisfying $W^\top W = I_N$. For any $f \in \mathbb{R}^N$, the rows of D defined by* (13) *form a frame system for $\mathbb{R}^N$. Furthermore, the rows of D form a tight frame system for $\mathbb{R}^N$ provided that $J(f)^\top J(f) = I_M$.*

Proof Notice that

$$\begin{aligned} S &= D(f)^\top D(f) \\ &= \frac{1}{2}W^\top (I_N + J(f)^\top J(f))W \\ &= \frac{1}{2}W^\top W + \frac{1}{2}W^\top J(f)^\top J(f)W \\ &= \frac{1}{2}I_N + \frac{1}{2}W^\top J(f)^\top J(f)W. \end{aligned}$$

Since the term $\frac{1}{2}W^\top J(f)^\top J(f)W$ is positive semi-definite, the matrix S is a positive definite matrix. Thus, the rows of $D(f)$ forms a frame for $\mathbb{R}^N$. Moreover, if $J(f)^\top J(f) = I_M$, then we have

$$S = \frac{1}{2}I_N + \frac{1}{2}W^\top J(f)^\top J(f)W = \frac{1}{2}I_N + \frac{1}{2}W^\top W = I_N.$$

Thus, the rows of $D(f)$ forms a tight frame for $\mathbb{R}^N$.

Proposition 1 states that the system formed by (13) is always a frame and is a tight frame if $J(f)$ is the analysis operator of some tight frame system. In the next, we give a detailed description on how to construct a non-local operator $J(f)$ whose rows form a tight frame for $\mathbb{R}^M$. First, the wavelet frame coefficients $Wf \in \mathbb{R}^M$ are grouped into P dis-adjoint sets $\{W_{\mathcal{G}_p} f\}_{p=1}^P$ with $\cup_{p=1}^P \mathcal{G}_p = \mathbb{Z}_M$ and $\mathcal{G}_{p_1} \cap \mathcal{G}_{p_2} = \emptyset$ if $p_1 \neq p_2$. Before introducing the ideal grouping strategy for partitioning wavelet coefficients, we first define the image region associated with each wavelet coefficient which will be used as the measurement for coefficient grouping.

Recall that in discrete multi-level wavelet decomposition, the k-th filter at the ℓ-th level, denoted by $a_k^{(\ell)}$, is defined as $a_k^{(\ell)} = a_0^{(\ell-1)} * a_k^{\downarrow_{2^{\ell-1}}}$, where $a_0^{(\ell-1)}$ and $a_k^{\downarrow_{2^{\ell-1}}}$ are given by (5) and (6). Then, the wavelet coefficients generated by the filter $a_k^{(\ell)}$ can be calculated by applying the filter $a_k^{\downarrow_{2^{\ell-1}}}$ on the image $f^{(\ell-1)} = f * a_0^{(\ell-1)}$, the $(\ell-1)$-th level low-pass output of the original image f in wavelet decomposition. Clearly, $f^{(\ell-1)}$ is f itself when $\ell = 1$ and is a smoothed version of f by the low-pass filter $a_0^{(\ell-1)}$ when $\ell > 1$. Suppose that all wavelet masks $\{a_k\}_{k=0}^m$ are

supported on $[-r,r]^2 \cap \mathbb{Z}^2$. Then, it can be seen that the filter $a_k^{\downarrow_{2^{\ell-1}}}$ is supported on $[-2^{\ell-1}r, 2^{\ell-1}r]^2 \cap \mathbb{Z}^2$, and the support of its non-zero elements is restricted on $[-2^{\ell-1}r, 2^{\ell-1}r]^2 \cap 2^{\ell-1}\mathbb{Z}^2$. In other words, the number of non-zero element of $a_k^{\downarrow_{2^{\ell-1}}}$ is the same as that of the corresponding mask a_k, i.e. $(2r+1)^2$.

For each wavelet coefficient $W_{j_0} f \in Wf$ which corresponds to the output of the filter $a_k^{(\ell)}$ with respect to image pixel located at i_0, we define its corresponding image region used for grouping by

$$\Omega(W_{j_0} f) = \{f^{(\ell-1)}(x) : x \in i_0 + [-2^{\ell-1}r, 2^{\ell-1}r]^2 \cap 2^{\ell-1}\mathbb{Z}^2\}. \tag{14}$$

In other words, the image region of each wavelet coefficient used for grouping is defined as the smoothed image $f^{(\ell-1)}$ sampled by the same way as its associated filter $a_k^{\downarrow_{2^{\ell-1}}}$. It can be seen that the image regions of all wavelet coefficients from different filters at different levels have the same size $(2r+1)\times(2r+1)$. Based on the definition (14), the ideal grouping strategy is summarized as follows.

Condition 2 (Ideal grouping strategy) *The wavelet coefficients $Wf \in \mathbb{R}^M$ are partitioned into P dis-adjoint sets $\{W_{\mathcal{G}_p} f\}_{p=1}^P$ such that $\cup_{p=1}^P \mathcal{G}_p = \mathbb{Z}_M$ and two wavelet coefficients $W_{j_1} f, W_{j_2} f$ belong to the same group if and only if*

(i) $W_{j_1} f$ and $W_{j_2} f$ are both generated by the same wavelet mask a_{k_0} with respect to either the same level or the different levels. In other words, they are generated by the filter sets $\{a_{k_0}^{(1)}, \ldots, a_{k_0}^{(L)}\}$ composed by the same mask a_{k_0} as (6) at all levels.
(ii) The mean square difference between two images regions $\Omega(W_{j_1} f)$ and $\Omega(W_{j_2} f)$ defined by (14) is bounded by a pre-defined threshold σ (very small).

Let $\{\mathcal{G}_p\}_{p=1}^P$ denote the index sets obtained via the ideal grouping strategy. Then $\{\mathcal{G}_p\}_{p=1}^P$ forms a dis-joint partition of $\mathbb{Z}_M$ that satisfies $\cup_p \mathcal{G}_p = \mathbb{Z}_M$ and $\mathcal{G}_{p_1} \cap \mathcal{G}_{p_2} = \emptyset$ if $p_1 \neq p_2$. Define the matrix $J \in \mathbb{R}^{M\times M}$ by

$$J(j,k) = \frac{1}{K_p} \begin{cases} -(K_p - 2), & j = k \text{ and } j \in \mathcal{G}_p; \\ 2, & j \neq k \text{ and } j,k \in \mathcal{G}_p; \\ 0, & \text{otherwise}, \end{cases} \tag{15}$$

where K_p denotes the cardinality of the set $\mathcal{G}_p$. The matrix J is defined such that the operation by J takes a weighted average of the input vector inside the same group. The weights are determined based on the following heuristic: large weight on the element itself, small weight equally on the other elements in the same group, and J forms an orthogonal matrix. Define the operator $D(f)$ by

$$D(f) = \frac{1}{\sqrt{2}} \begin{pmatrix} I \\ J \end{pmatrix} W = \frac{1}{\sqrt{2}} \begin{pmatrix} W \\ JW \end{pmatrix}. \tag{16}$$

Then such an operator $D(f)$ is indeed an analysis operator of a tight frame for $\mathbb{R}^N$.

Proposition 3 *Let $W \in \mathbb{R}^{M\times N}$ be a tight frame system satisfying $W^\top W = I_N$. For any $f \in \mathbb{R}^N$, the rows of $D(f)$ defined by (16) forms a tight frame for $\mathbb{R}^N$.*

Proof Let $\mathcal{G} \in \mathbb{R}^M$ denote the vector of all integer indexes formed by sequentially concatenating the index sets $\{\mathcal{G}_p\}_{p=1}^P$ obtained via the ideal grouping strategy:

$$\mathcal{G} = (\mathcal{G}_1^\top, \ldots, \mathcal{G}_P^\top)^\top.$$

By the definition of $\{\mathcal{G}_p\}_p$, the vector $\mathcal{G}$ is a permutation of the vector $(1, \ldots, M)^\top$. Let Q denote the corresponding permutation matrix such that $Q\mathcal{G} = [1, 2, \ldots, M]^\top$. Then, we have $Q^\top Q = I_M$, and by the definition (15),

$$Q^\top J = \mathrm{diag}(J_1, J_2, \ldots, J_P),$$

where $\mathrm{diag}(J_1, J_2, \ldots, J_P)$ denotes the block-wise diagonal matrix formed by the sub-matrices $\{J_p\}_p$, and $J_p \in \mathbb{R}^{K_p \times K_p}$ is given by

$$J_p = \frac{1}{K_p} \begin{pmatrix} 2-K_p & 2 & \cdots & 2 & 2 \\ 2 & 2-K_p & \cdots & 2 & 2 \\ 2 & 2 & \ddots & 2 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2 & 2 & \cdots & 2 & 2-K_p \end{pmatrix}.$$

A direct calculation leads to ${J_p}^\top J_p = I_{K_p}$. Thus,

$$\begin{aligned} J(f)^\top J(f) &= \mathrm{diag}(J_1, J_2, \ldots, J_P)^\top Q^\top Q \mathrm{diag}(J_1, J_2, \ldots, J_P) \\ &= \mathrm{diag}(J_1^\top J_1, \ldots, J_P^\top J_P) \\ &= I_M. \end{aligned}$$

By Proposition 1, the rows of $D(f)$ forms a tight frame for $\mathbb{R}^N$.

It can be seen that the non-zero elements of the operator $J(f)$ sparsely spread out over the whole matrix. As a result, the elements of the tight frame, the rows of $J(f)W$, have a non-local support. Indeed, as we show in the following, the outcome $J(f)Wf$ is a non-local version of wavelet tight frame coefficient vector Wf. By the definition of $\mathcal{G}_p$, we have that

$$(Wf)(j) = (Wf)(k), \text{for any } j, k \in \mathcal{G}_\ell$$

which is equivalent to

$$(Wf)(j) = \frac{1}{K_\ell}\Big((\sum_{k \neq j, k \in \mathcal{G}_\ell} 2(Wf)(k)) - (K_\ell - 2)Wf(j) \Big), \quad \forall\, j \in \mathcal{G}_\ell, \tag{17}$$

or in the matrix form:

$$Wf = J(f)Wf.$$

Thus, when we use $D(f)$ defined by (16) in the regularization model (12), the regularization functional includes two components:

$$\begin{aligned} \|\mathrm{diag}(\lambda) D(f) f\|_1 &= \| \begin{pmatrix} \mathrm{diag}(\lambda_1) \\ \mathrm{diag}(\lambda_2) J(f) \end{pmatrix} Wf \|_1 \\ &= \|\mathrm{diag}(\lambda_1) Wf\|_1 + \|\mathrm{diag}(\lambda_2) J(f) Wf\|_1. \end{aligned} \tag{18}$$

The first component is the local version of the weighted ℓ_1 norm of wavelet tight frame coefficients and the second component is the non-local version of the weighted ℓ_1 norm of wavelet tight frame coefficients. Compared to the widely used regularization functional $\|\mathrm{diag}(\lambda)Wf\|_1$, the new functional $\|\mathrm{diag}(\lambda)D(f)f\|_1$ not only enforces the same local sparsity prior of the image f in wavelet tight frame domain, but also enforces two additional non-local constraints on the self-recursions of wavelet frame coefficients: one in spatial domain and the other in scale domain.

In practice, there are a few issues to be addressed when directly using the nonlocal tight frame $D(f)$ constructed by (16) in image restoration. The first is we usually have a rough estimate of f in hand in image restoration which may contain noticeable errors in the early stage. The second is the imperfection of self-recursions of image structures in practice, as they are usually not exactly repeating themselves but rather they are highly similar. The last is the computational efficiency as a rigorous treatment on grouping requires an overwhelming amount of computational cost. Thus, we propose a robust version of $D(f)$ to address these issues. The modified system is still of the form (16) but with a slightly different non-local operator. It is noted that the robust version of $D(f)$ is a frame, not a tight frame.

The first modification for the robustness is on the construction of $\{\mathcal{G}_p\}_{p=1}^P$. Instead of grouping only wavelet coefficients corresponding to the same image structure, we group the wavelet coefficients corresponding to image structures with high similarity. Then the weights are introduced into the definition of $J(f)$ to encode the similarity degree of wavelet frame coefficients. Let $\{\mathcal{H}_r\}_{r=1}^R$ denote the groups formed by the new strategy. The modified non-local operator, denoted by $\widetilde{J}$, is then defined as follows,

$$\widetilde{J}(j,k) = \frac{1}{\omega(j,\mathcal{H}_r)+1}\begin{cases} -(\omega(j,\mathcal{H}_r)-1), & j=k; \\ 2\omega(j,k), & j\neq k \text{ and } j,k\in\mathcal{H}_r; \\ 0, & \text{otherwise}, \end{cases} \tag{19}$$

where $0 < \omega(j,k) \leq 1$ is the weight defined by some similarity measurement between two wavelet tight frame coefficients, and $\omega(j,\mathcal{H}_\ell) = \sum_{k\neq j, k\in\mathcal{H}_\ell}\omega(j,k)$. It can be seen that the definition of (19) is degenerated to (15) of the ideal case where all weights are the same.

In summary, we propose a modified version of the operator $D(f)$ for better performance when solving practical image restoration problems:

$$D(f) = \frac{1}{\sqrt{2}}\begin{pmatrix} I \\ \widetilde{J}(f) \end{pmatrix} W \tag{20}$$

where $\widetilde{J}(f)$ is defined by (19) and W is the analysis operator of some wavelet tight frame system. By proposition 1, the new operator $D(f)$ of the form (20) is the analysis operator of a frame system for $\mathbb{R}^N$.

3.2 Regularization model for image recovery

For image restoration, the non-local adaptive tight frame proposed in the previous section cannot be pre-constructed as the true image f is not available. Thus, we

propose the following minimization model for solving the ill-posed linear system (1):

$$\text{argmin}_{f\in\mathbb{R}^N}\frac{1}{2}\|Af-g\|_2^2+\|\text{diag}(\lambda)D(f)f\|_1, \tag{21}$$

where $D(f)$ is defined by (20). We take an iterative scheme to alternatively update the estimation of f and the operator $D(f)$. See Algorithm 1 for the outline of the iterative scheme. There are two non-trivial steps in Algorithm 1, one is using the current estimate $f^{(k)}$ of the true image to construct the index sets $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$ and the other is to re-estimate the true image by solving the minimization model (21) with $D(f) := D^{(k)}$. The construction of $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$ in first step requires solving a discrete optimization problem. Regarding the minimization problem (21), many efficient numerical solvers have been developed for solving such type of minimization problems in recent years, e.g, the split Bregman iteration [32,33].

Algorithm 1 Alternative iteration scheme for solving (21)

Input: the degraded image g
Output: the recovered image f
Main procedure:

1. $f^{(0)} := g$.
2. for $k = 0, 1, \ldots, K-1$,
   (a) constructing the index sets of groups $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$ using $f^{(k)}$;
   (b) synthesizing $D^{(k)}$ defined by (20) and (19) using $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$;
   (c) $f^{(k+1)} := \text{argmin}_{f\in\mathbb{R}^N}\frac{1}{2}\|Af-g\|_2^2+\|\text{diag}(\lambda)D^{(k)}f\|_1$.
3. $f := f^{(K)}$.

In the computation, a rigorous construction of $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$ is a very time consuming process. Certain simplifications are needed such that the construction provides a reasonable good grouping within acceptable running time. In our implementation, for computational efficiency, all index sets $\{\mathcal{H}_r\}_{r=1}^R$ have the same size S and for each framelet coefficient, only the first $S-1$ framelet coefficients in its neighborhood with the highest similarity are grouped together. Then, the similarity measure between two framelet coefficients c_i and c_j from the same group $\mathcal{H}_r^{(k)}$ is defined as

$$\omega(c_i,c_j)=\begin{cases}\beta e^{-d(c_i,c_j)/h}, & \text{if } c_i, c_j \text{ relate to the same filter;}\\ (1-\beta)e^{-d(c_i,c_j)/h}, & \text{otherwise,}\end{cases} \tag{22}$$

where $0<\beta<1$ determines percentage of inner-scale and cross-scale nonlocal constraints, h is the constant which is set to 0.3 in our implementation, and $d(c_i,c_j)$ is defined as the square difference summation of two image regions associated with the two coefficients as defined in (14).

The error in the initial estimates $f^{(k)}$ will also have a negative impact on the performance of the deconvolution process by (21). The negative impact mainly comes from the relatively low quality of the non-local operator $\widetilde{J}(f)$ in the frame system $D^{(k)}$. Thus, we propose to update the regularization parameter vector $\lambda^{(k)}$ during each iteration according to the quality of the operator $D^{(k)}$. In our

implementation, we propose the following empirical formulation of setting the regularization parameter vector $\lambda^{(k)} \in \mathbb{R}^{2M}$:

$$\lambda^{(k)} = \begin{pmatrix} \lambda_1^{(k)} \\ \lambda_2^{(k)} \end{pmatrix}, \tag{23}$$

where

$$\begin{cases} \lambda_1^{(k)}(j) = c_0\lambda_0\sigma_n^2(\eta + \max(0, \frac{1}{S}\sum\limits_{i,j\in H_r}|Wf^{(k)}(i)|^2 - \sigma_n^2))^{-1/2}, \\ \lambda_2^{(k)}(j) = (1-c_0)\lambda_0\sigma_n^2(\eta + \max(0, \mathrm{var}(\{Wf^{(k)}(i) : i,j \in H_r\}) - \sigma_n^2))^{-1/2}. \end{cases} \tag{24}$$

In (24), the variable σ_n denotes the standard deviation of image noise which could be estimated in many image restoration tasks by some statistical approach; see e.g. [34]. The operator $\mathrm{var}(\cdot)$ denotes the variance estimator. There are three constants in (24): λ_0, c_0, η. The constant λ_0 is the regularization value; the constant $c_o \in [0,1]$ determines the percentage of local and non-local constraints in the overall regularization; the constant η is just for numerical stability. Once the regularization parameter vector $\lambda^{(k)}$ is set, the minimization (21) can be solved by split Bregman iteration. A detailed description for the proposed regularization method is summarized in Algorithm 2. The subproblem $u^{(p)}$ in Step 2 (d) is solved by the conjugate gradient method with 20 iterations in our implementation.

Algorithm 2 Regularization method using non-local tight frames

Input: the observed image g
Output: the recovered image f
Main procedure:

1. $f^{(0)} := g$.
2. For $k = 1, ..., K$,
   (a) constructing the index sets of groups $\{\mathcal{H}_r^{(k)}\}_{r=1}^R$ using $f^{(k)}$.
   (b) synthesizing $D^{(k)}$ using (20).
   (c) $d^{(0)} = b^{(0)} := \mathbf{0}$ and $D := \widetilde{D}^{(K)}$.
   (d) For $p = 1, ..., P$,
      i. synthesizing $\lambda^{(k)}$ using (23) and (24).
      ii. defining

$$\begin{cases} u^{(p)} := (A^TA + \mu D^TD)^{-1}(A^Tg + \mu D^T(d^{(p-1)} - b^{(p-1)})); \\ d^{(p)} := T_\lambda(Du^{(p)} + b^{(p-1)}); \\ b^{(p)} := b^{(p-1)} + \delta(Du^{(p)} - d^{(p)}), \end{cases},$$

      where μ and δ are two parameters used in split Bregman iteration, and T_λ is the soft-thresholding operator defined by

$$T_\lambda(u) = \mathrm{sign}(u)\max(|u| - \lambda, 0).$$

   (e) $f^{(k)} := u^{(P)}$.
3. $f := f^{(K)}$.

Fig. 1 Eight tested images

4 Experiments and discussions

Algorithm 2 can be used to solve general image restoration problems. In this section, we will mainly focus on the evaluation of Algorithm 2 in the application of image deconvolution. The applications of Algorithm 2 in image de-noising and image in-painting will be demonstrated only with a few examples.

Let $\mathbf{f}$ and $\mathbf{g}$ denote the true image and the observed image, and let $f \in \mathbb{R}^N$ and $g \in \mathbb{R}^N$ denote their vectorized versions, where $n \in \mathbb{R}^N$ denotes image noise. For image deconvolution, the matrix A in the linear system (1) becomes the convolution operator, i.e., the blurring process can be modeled as a convolution process such that

$$\mathbf{g} = \mathbf{p} * \mathbf{f} + \mathbf{n}, \tag{25}$$

where $\mathbf{p}$ denotes the blur kernel and '$*$' denotes the 2D discrete convolution operator. The system (25) can be re-written in the matrix form:

$$g = Af + n, \tag{26}$$

where A is a block-wise Toeplitz matrix under Neumann boundary extension or a block-wise circulant matrix under periodic boundary extension.

4.1 Experiments and demonstrations

For the experiments on image deconvolution, we run Algorithm 2 on eight images as shown in Figure 1. The experimental setup is as follows. Each tested image is first convoluted by a blur kernel, followed by the addition of white Gaussian noise with different standard deviations. Four different types of blur kernels are tested in the experiments, including (a) a disk kernel of radius 3 pixels; (b) a linear motion kernel of length 15 pixels and orientation 30°; (c) a Gaussian kernel of size 25×25 pixels and standard deviation $8/5$; and (d) a box kernel of size 9×9 pixels. The mathematical expression of a Gaussian blur kernel is $p(x, y) =$

$(2\pi\sigma^2)^{-1}\exp(-(x^2+y^2)/(2\sigma^2))$. The other three blur kernels can be expressed as

$$\mathbf{p}(x) = \frac{1}{\#\Omega}\begin{cases} 1, x \in \Omega, \\ 0, \text{otherwise}, \end{cases}$$

where Ω denotes the support of the kernel, which is a disk, a box and a line segment respectively; and $\#\Omega$ denotes the cardinality of Ω. We use the peak signal to noise ratio (PSNR) as the quantitative measure of image quality which is defined by

$$\text{PSNR}(f,\hat{f}) = -20\log_{10}\frac{||f-\hat{f}||_2^2}{255N},$$

where f denotes the true image, $\hat{f}$ denotes the restored image and N denotes the number of image pixels.

The experimental results from Algorithm 2 are compared to that from three related existing image restoration approaches. The first one is the framelet based image restoration approach [33] which used the same minimization model (21) to recover the true image. The main difference between the framelet based method and Algorithm 2 lies in the choice of the wavelet frame system D in the model (21). The former used the linear spline wavelet frame system and the later used the non-local wavelet frame system of the form (20). The second method for comparison is the non-local TV based approach [22] which introduced a non-local mean operator into the TV-based image regularization. The third one is the patch-based BM3DDEB approach introduced in [25], the generalization of the BM3D method to image deconvolution. The BM3DDEB method is built on the concept of collaborative filtering on the 3D array of matched image patches. The results from these three methods for comparison are all generated by using the codes with recommended parameter settings from the original authors.

The same parameters of Algorithm 2 are used for all eight input images during the experiments. As we observed in the experiments, only one more round of the refinement for patch grouping is sufficient to obtain satisfactory results. Thus, the iteration number K of the outer loop is set to 2. The iteration number P of the inner loops is set to 30. The two parameters μ and δ for the spilt Bregman iteration are set to 1/10 and 1 respectively. The parameter β in (22) is set to 1/20. The parameters of the adaptive regularization vector defined in (24) are set to $\eta = 3/10$, $\lambda = 2\times 10^{-3}$ and $c_0 = 6/10$. The fact that $c_0 > 0.5$ implies that the regularization in (21) weights more on the non-local wavelet frame coefficients. On a PC workstation with a 6 core INTEL Xeon CPU (2.4 Ghz) and 32 GB memory, the MATLAB implementation of Algorithm 2 takes about 9 minutes for an image with size 256×256 in total, including 2 minutes for patch grouping.

Table 1 and 2 summarized the PSNR values of the results de-blurred by four methods, with respect to different configurations on blur kernels and noise levels. It can be seen that Algorithm 2 outperformed the framelet method and the non-local TV method by a large margin in most test images. Also it outperformed the BM3DDEB method in most test images, but it did not perform as well as the BM3DDEB method on the image "Barbara512". The main reason is that the BM3DDEB method enforces the sparsity constraint in DCT transform, which is very suitable for representing the stripe textures that are prevailing in the image "Barbara512". For the images dominated by other types of texture or dominated by cartoon-type regions, Algorithm 2 performed noticeably better. The reason

Image	Kernel	Local framelet	Nonlocal TV	BM3DDEB	Algorithm 2
peppers256	disk	28.43	28.08	30.35	**31.12**
	motion	28.35	27.25	29.53	**30.17**
	gaussian	26.73	26.75	27.19	**29.84**
	box	28.46	27.60	28.55	**29.48**
goldhill256	disk	27.66	27.63	27.87	**28.16**
	motion	27.66	26.97	27.41	**28.19**
	gaussian	27.10	27.28	27.38	**27.73**
	box	26.87	26.57	26.89	**27.33**
boat256	disk	26.88	27.06	27.13	**27.87**
	motion	26.75	26.59	26.92	**27.63**
	gaussian	26.31	26.72	26.57	**26.90**
	box	25.64	25.72	25.93	**26.30**
camera256	disk	27.45	28.06	28.54	**29.11**
	motion	27.63	27.57	28.29	**29.02**
	gaussian	26.39	27.00	27.08	**27.34**
	box	25.77	26.12	26.61	**26.85**
bridge256	disk	25.74	25.67	25.75	**26.13**
	motion	25.67	24.93	25.43	**25.78**
	gaussian	25.30	25.49	25.51	**25.81**
	box	24.61	24.30	24.67	**24.85**
house256	disk	32.98	32.51	33.54	**34.52**
	motion	32.10	31.13	32.72	**33.79**
	gaussian	31.80	32.41	32.41	**33.23**
	box	31.79	30.71	32.32	**33.11**
Barbara512	disk	25.47	25.76	**27.63**	26.33
	motion	25.49	25.48	**27.99**	26.86
	gaussian	24.58	24.70	**25.05**	24.71
	box	24.30	24.31	**25.01**	24.61
Lena512	disk	32.91	31.82	33.55	**33.81**
	motion	32.03	30.46	32.63	**33.04**
	gaussian	33.05	32.90	33.50	**33.98**
	box	30.68	30.01	31.06	**31.47**

Table 1 Comparison of the PSNR values (dB) of the results by four methods, with respect to the noise level $\sigma = 2$.

is that the self-recursive property of image structures is very weak on these two images which results in reliable patch stacks in the BM3DDEB method. Contrary to the BM3DDEB method, the non-local wavelet frame system used in Algorithm 2 allows the simultaneous usage of both the local prior in wavelet frame domain and the non-local self-recursion prior of wavelet frame coefficients. Together with the regularization parameters adaptive to the accuracy of each prior, Algorithm 2 performs consistently on a wide range of images, including both image of rich textures and images of less textures. The advantage of Algorithm 2 over other methods in terms of the PSNR value is also consistent with the improvement of the visual quality. See Figures 2, Figure 3 and Figure 4 for the visual comparison of the results on a few tested images.

Algorithm 2 can also be applied to solve many other image restoration problems with only small modifications. In the end of the experimental section, we demonstrated some examples of applying Algorithm 2 to image de-noising and image in-painting. For image de-noising, the matrix A in the linear system (1) becomes the identical matrix and for image in-painting, the matrix A in the linear system (1) becomes a diagonal matrix with diagonal element being 1 if the corresponding

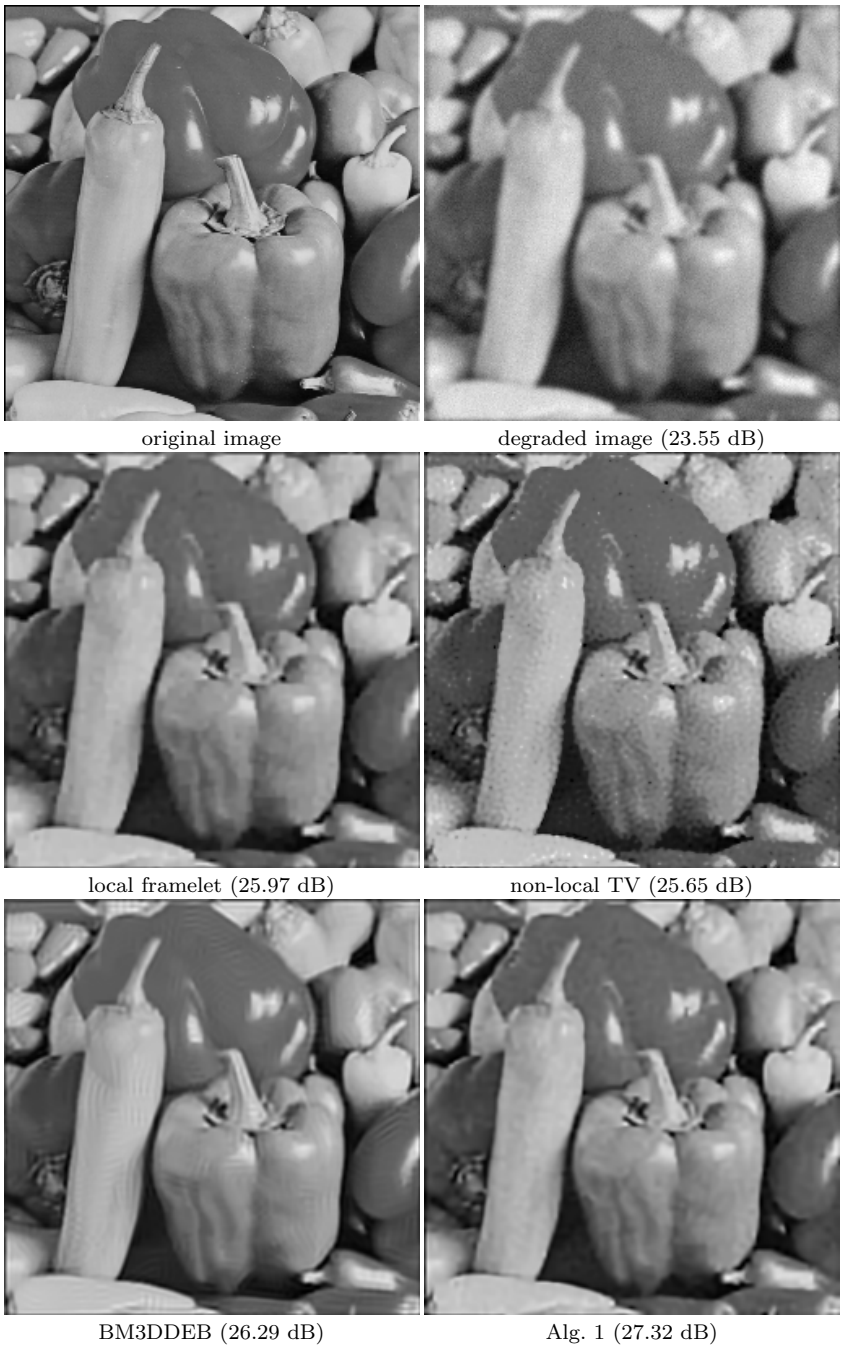

Fig. 2 Visual comparison of de-blurred results for the image "peppers256". The true image is degraded by the 25×25 Gaussian kernel with standard deviation 1.6 and the Gaussian noise with standard deviation $\sigma = 5$.

Fig. 3 Visual comparison of de-blurred results for the image "cameraman256". The true image is degraded by the 9×9 box kernel and the Gaussian noise with standard deviation $\sigma = 5$.

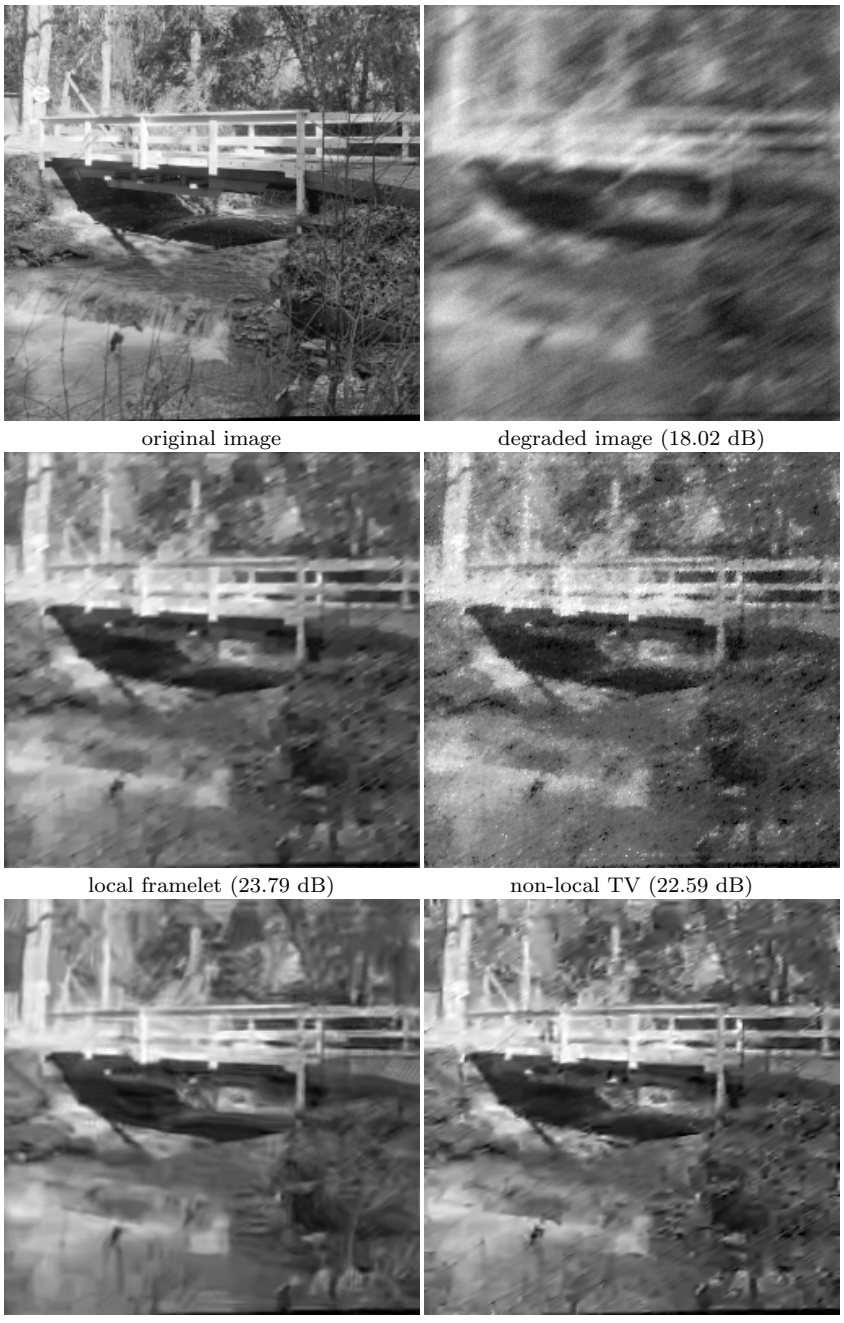

Fig. 4 Visual comparison of de-blurred results for the image "bridge256". The true image is degraded by the 9×15 motion kernel and the Gaussian noise with standard deviation $\sigma = 5$.

Image	Kernel	Local framelet	Nonlocal TV	BM3DDEB	Algorithm 2
peppers256	disk	26.94	25.22	28.08	**28.37**
	motion	25.41	24.20	26.60	**27.46**
	gaussian	25.97	25.65	26.29	**27.32**
	box	25.84	25.45	26.67	**27.71**
goldhill256	disk	26.55	25.43	26.65	**26.75**
	motion	25.93	24.44	25.95	**26.27**
	gaussian	26.33	26.29	26.59	**26.88**
	box	25.40	24.87	25.69	**26.00**
boat256	disk	25.23	24.91	25.71	**25.87**
	motion	24.61	23.88	25.06	**25.44**
	gaussian	25.40	25.58	25.64	**25.85**
	box	24.19	24.20	24.56	**24.88**
camera256	disk	25.43	25.43	26.50	**26.64**
	motion	24.97	24.33	25.82	**26.43**
	gaussian	25.44	25.91	26.02	**26.15**
	box	24.16	24.51	24.85	**25.32**
bridge256	disk	24.33	23.51	24.55	**24.78**
	motion	23.79	22.59	23.87	**24.22**
	gaussian	24.51	24.38	24.63	**24.85**
	box	23.18	22.79	23.49	**23.75**
house256	disk	30.75	29.50	**32.02**	31.88
	motion	29.78	27.28	30.87	**31.07**
	gaussian	30.72	30.54	31.19	**31.64**
	box	29.48	28.66	30.34	**30.86**
Barbara512	disk	24.28	24.30	**25.28**	24.79
	motion	24.10	23.73	**25.01**	24.48
	gaussian	24.27	24.15	**24.43**	24.40
	box	23.69	23.54	**24.01**	24.00
Lena512	disk	31.13	28.95	31.72	**31.82**
	motion	29.77	27.45	30.30	**30.64**
	gaussian	31.53	30.63	32.15	**32.24**
	box	29.17	28.18	29.35	**29.96**

Table 2 Comparison of the PSNR values (dB) of the results from the four methods, with respect to the noise level $\sigma = 5$.

pixel is known or 0 otherwise. In this paper, we only provide a few examples. See Figure 5 for the visual illustration of the result de-noised by the model (9) using 3-level linear spline wavelet frame and the one using the non-local wavelet frame (20). See Figure 6 and Figure 7 for the comparisons of the in-painted results by the linear spline wavelet approach and the non-local wavelet frame approach. It is seen that the non-local wavelet tight frame (20) performed better than the linear spline framelet on the tested images. We also compare our approach with the recent non-local method from [35]. The visual results can be refereed to Fig. 8. Obviously, our approach outperformed the non-local method from [35], which demonstrates the power of the combination of the local and non-local priors.

4.2 Summary

The performance of a regularization-based image restoration method largely depends on not only the accuracy but also the strength of the assumptions used for regularizing the true image. In recent years, the wavelet tight frame based

noisy image (22.10 dB) linear spline framelet (31.61 dB) Alg. 2 (32.78 dB)

Fig. 5 Visual comparison of de-noised results for the image "Lena512". The true image is degraded by Gaussian noise with standard deviation 20.

damaged image (40% lost) linear spline framelet (28.46 dB) Alg. 2 (30.05 dB)

Fig. 6 Visual comparison of in-painted results for the image "boat256". The forty percents of image pixels of the true image are randomly missing.

damaged image linear spline framelet (34.38 dB) Alg. 2 (35.40 dB)

Fig. 7 Visual comparison of in-painted results for the image "Lena512". Some image pixels of the true image are damaged by the scratches and texts.

regularization plays an active role in the newest development of powerful image restoration methods, which seek the solution that minimizes the ℓ_1 norm of the associated wavelet tight frame coefficients. The wavelet tight frame approach assumes that the image of interest is likely to have sparse approximation in the wavelet tight frame domain, which essentially exploits the sparse nature of local image intensity variations. However, in existing wavelet tight frame systems, there is no measurement on the self-similarity relationship of image structure in spatial domain, one often seen phenomena in natural images of complex textures. By ex-

damaged image Result from [35] (28.16 dB) Alg. 2 (30.10 dB)

Fig. 8 Visual comparison of the in-painted results of our method and [35]. The fifty percents of image pixels of the true image are randomly missing.

ploiting such non-local self-recurrence , the patch-based methods like BM3D or the non-local means methods demonstrated the promising performance in various image restoration tasks.

Motivated by the idea behind these non-local approaches, we developed in this paper a scheme of constructing non-local wavelet frame and tight frame that are adaptive to the input image. The constructed adaptive non-local wavelet frames are composed of two types of wavelet frame systems: one local version and one nonlocal version of some existing wavelet frame system. The proposed wavelet frame system not only measures the local variations on image intensity, but also measures two self-recurrence properties of image structures: one in image domain and one in scale domain owing to the multi-resolution nature of wavelet system. As such, the ℓ_1 norm based regularization under the proposed non-local wavelet frame can simultaneously exploit all these three image priors. The experiments on image de-convolution showed the advantages of our approaches over some related approaches.

References

1. I. Daubechies. *Ten lectures on wavelets.* CBMS-NSF Lecture Notes, SIAM, first edition, 1992.
2. Stéphane Mallat. *A wavelet tour of signal processing.* Academic Press Inc., San Diego, CA, third edition, 2009.
3. A. Ron and Z. Shen. Affine system in $L_2(R^d)$: the analysis of the analysis operator. *J. of Func. Anal.*, 148, 1997.
4. I. Daubechies, B. Han, A. Ron, and Z. Shen. Framelets: MRA-based constructions of wavelet frames. *Appl. Comput. Harmon. Anal.*, 14:1–46, 2003.
5. Z. Shen. Wavelet frames and image restorations. In *Proc. ICM*, Hyderabad, India, 2010.
6. D. Donoho. De-noising by soft thresholding. *IEEE Trans. Info. Theory*, 41(3):613–627, 1995.
7. E. Candes and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise-C^2 singularities. *Comm. Pure Appl. Math*, 57:219–266, 2002.
8. J. Starck, E. Candes, and D. Donoho. The curvelet transform for image denoisi. *IEEE Trans. Image Proc.*, 11(6):670–684, 2002.

9. J. Cai, H. Ji, C. Liu, and Z. Shen. High-quality curvelet-based motion deblurring using an image pair. In *CVPR*, 2009.
10. S. Mallat and E. Lepennec. Sparse geometric image representation with bandelets. *IEEE Trans. Image Proc.*, 14:423–438, 2005.
11. E. Pennec and S. Mallat. Bandlet image approximation and compression. *SIAM Multiscale Model. Simul.*, 4(3):992–1039, 2005.
12. J.-F. Cai, R. Chan, and Z. Shen. A framelet-based image inpainting algorithm. *Appl. Comput. Harmon. Anal.*, 24:131–149, 2008.
13. J. Cai, H. Ji, C. Liu, and Z. Shen. Framelet based blind image deblurring from a single image. *IEEE Trans. Image Proc.*, 21(2):562–572, 2012.
14. B. Dong, H. Ji, J. Li, and Z. Shen. Wavelet frame based blind image inpainting. *Appl. Comput. Harmon. Anal.*, 32(2):268–279, 2012.
15. L. Rudin, S. Osher, and E. Fatemi. Nonllienar total variation based noise removal algorithms. *Phys. D.*, 60:259–268, 1992.
16. J. Cai, B. Dong, Z. Shen, and S. Osher. Image restoration: total variation; wavelet frames; and beyond. *J. Amer. Math. Soc.*, 25(4):1033–1089, 2012.
17. M. Elad and M. Ahron. Image denoising via sparse and redundant representations over learned distionaries. *IEEE Trans. Image Proc.*, 54(12):3736–3745, 2006.
18. J. Cai, S Huang, Z. Shen, and G. Ye. Data-driven tight frame construction and image denoising. Technical report, UCLA-CAM report 12-40, 2012.
19. A. Buades, B. Coll, and J.M. Morel. A non-local algorithm for image denoising. In *CVPR*, 2005.
20. G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *SIAM Multiscale Model. Simul.*, 7(3):1005–1028, 2008.
21. G. Peyre, S. Bougleux, and L. Cohen. Non-local regularization of inverse problems. In *ECCV*, 2008.
22. Y. Lou, X. Zhang, S. Osher, and A. Bertozzi. Image recovery via nonlocal operators. *Journal of Scientific Computing*, 42(2):185–197, 2010.
23. X. Zhang, M. Burger, X. Bresson, and S. Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM J. Imaging Sci.*, 3(3):253–276, 2010.
24. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans. Image Proc.*, 16(8):2080–2095, 2007.
25. K. Dabov, A. Foi, K. Katkovnik, and K. E. Image restoration by sparse 3d transform-domain collaborative filtering. *SPIE Electronic Imaging*, 6812, 2008.
26. A. Danielyan, V. Katkovnik, and K. Egiazarian. BM3D frames and variational image deblurring. *IEEE Trans. Image Proc.*, 21(4):1715–1728, 2012.
27. J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *ICCV*, 2011.
28. B. Dong and Z. Shen. MRA based wavelet frames and applications. *IAS Lecture Notes Series, Park City Mathematics Institute*, 2010.
29. J.-F. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring from a single image using sparse approximation. In *CVPR*, 2009.
30. R. Chan., S. Riemenschneider, L. Shen, and Z. Shen. Tight frame: an efficient way for high-resolution image reconstruction. *Appl. Comput. Harmon. Anal.*, 17:91–115, 2004.
31. M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse Problems*, 23(3):947–968, 2007.
32. T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM J. Img. Sci.*, 2(2):323–343, 2009.
33. J.F. Cai, S. Osher, and Z. Shen. Split bregman methods and frame based image restoration. *SIAM Multiscale Model. Simul.*, 8(2):337–369, 2009.
34. D. Donoho and I. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.*, 90:1200–1224, 1995.
35. Jianping Shi, Xiang Ren, Guang Dai, Jingdong Wang, and Zhihua Zhang. A non-convex relaxation approach to sparse dictionary learning. In *CVPR*. IEEE, 2011.