Data-driven tight frame construction and image denoising

Jian-Feng Cai[a], Sibin Huang[b], Hui Ji[*,b], Zuowei Shen[b], Gui-Bo Ye[c]

[a] *Department of Mathematics, University of Iowa, Iowa City, Iowa 52242*
[b] *Department of Mathematics, National University of Singapore, Singapore 117543*
[c] *Department of Mathematics, University of California, Irvine, CA 92617*

Abstract

The regularization methods for image restoration using the ℓ_1 norm of the coefficients of the underlying image under some system assume that the image has a good sparse approximation under the given system. Such a system can be a basis, a frame or a general over-complete dictionary. One widely used system in image restoration is wavelet tight frame. There have been enduring efforts on seeking wavelet type of tight frames under which certain class of functions or images can have a good sparse approximation. However, the structure of images varies greatly in practice and a system working well for one type of images may not work for another. This paper presents a method that derives discrete tight frame system from the input image itself to provide a better sparse approximation to the input image. Such an adaptive tight frame construction scheme is applied on image denoising by constructing a tight frame tailor down to the given noisy data. The experiments showed the proposed approach performs better in image denoising than those wavelet tight frames designed for a class of images. Moreover, by ensuring the system derived from our approach is always a tight frame, our approach also runs much faster than some other adaptive over-complete dictionary based approaches with comparable PSNR performance.

Key words: tight frame, image de-noising, wavelet thresholding, sparse approximation

1. Introduction

In the past decades, sparse approximation has been playing a fundamental role in many signal processing areas, such as compression, data analysis and signal restoration. Sparse approximation is about keeping most information

*Corresponding author

Email addresses: jianfeng-cai@uiowa.edu (Jian-Feng Cai), mathsb@nus.edu.sg (Sibin Huang), matjh@nus.edu.sg (Hui Ji), matzuows@nus.edu.sg (Zuowei Shen), yeg@uci.edu (Gui-Bo Ye)

of the given data with a linear combination of a small number of atoms of some system. Among many different systems, orthonormal wavelet basis [1] has been very successful in signal processing as it can approximate piece-wise smooth 1D signals very efficiently with only few non-zero wavelet coefficients. In recent years, over-complete system has become more and more recognized and used in signal processing. Over-complete system has several advantages over orthonormal basis for sparsely approximating signals, as signals are more likely to have a good sparse approximation under a redundant system. Also, it offers more flexibilitis and conveniences in the filter design. One representative over-complete system is the so-called *wavelet tight frame* ([2, 3]) which is now widespread in many signal processing tasks. Wavelet tight frame sacrifices the orthonormality and linear independence of orthonormal basis while still enjoying the same efficient decomposition and reconstruction schemes as orthonormal wavelet basis.

Most wavelets used in image processing are separable wavelet bases defined from the tensor product of the 1D wavelet bases. Despite their successes in 1D signal processing, tensor wavelets are much less efficient when approximating images since tensor wavelets mostly focus the horizontal and vertical discontinuities of images. When the discontinuities of the target image has complex geometric properties, the sparsity of a good approximation under tensor wavelets is not satisfactory. In recent years, many tight frames have been proposed to more efficiently represent nature images, including *ridgelet* [4], *curvelet* [5, 6], *bandlet* [7], *shearlet* [8] and many others. However, the efficiency of these redundant systems heavily relies on certain functional assumptions of nature images, e.g. isolated objects with C^2 singularity assumed by curvelet. Such an assumption is applicable to cartoon-type images but not to textured images. Nature images vary greatly in terms of the geometrical structure, and often they contain a significant percentage of irregular textures with fractal structures. A tight frame designed for efficiently representing one type of continuum may not provide a good sparse approximation to the input image. A better approach is to develop a tight frame system that is specifically optimized for the given image. In other words, the design of a tight frame system should be driven by the input data in order to achieve great performance in terms of sparse approximation.

The concept of "adaptivity" has been explored in recent years by the so-called learning approaches (e.g., [9, 10, 11, 12, 13]). The learning approach learns an over-complete dictionary from the input image itself to achieve better sparsity of the input image over the learned dictionary. The basic idea of most existing approaches is first partitioning images into small image patches and then finding a set of atoms of the dictionary such that each image patch can be approximated by a sparse linear combination of atoms in the dictionary. The adaptively learned over-complete dictionaries by these approaches are very effective on sparsely approximating nature images with rich textures. As a result, these adaptive over-complete dictionary based approaches tend to outperform the sparsity-based wavelet thresholding methods in image denoising. Despite the success of these adaptively learned approaches, the over-complete dictionaries constructed by these approaches lack several properties desired for image

restoration. One is the so-called *perfect reconstruction property* which ensures that the given signal can be perfectly represented by its canonical expansion in a manner similar to orthonormal bases. Also, finding an optimal over-complete system often leads to a severely under-constrained ill-posed problem owing to the redundancy of the over-complete system. It remains a challenging task to develop fast and stable numerical methods for estimating an optimal over-complete system.

In this paper, we aim at developing an new approach to construct discrete tight frame that is adaptive to the input image. The adaptively learned tight frame in our proposed approach is more likely to give a highly sparse approximation of the input image than existing wavelet tight frames. Different from general over-complete dictionary, tight frame satisfies the perfect reconstruction property which is appealing to many image restoration tasks. Also, the sparsity of canonical frame coefficients of an image is closely related the regularity of the image, which are assumed by many image restoration approaches to obtain the results with less artifacts (see more details in [14]). Moreover, as we will show later, the minimization problems arising in the construction of tight frame are much better conditioned than that of generic over-complete dictionary, owing to the *unitary extension principle* [14] satisfied by wavelet tight frame. Thus, by considering a class of tight frames with certain special properties, a very fast numerical method is available to construct data-driven tight frame.

To illustrate the benefit of adaptively constructed tight frame from the data itself, we derived an adaptive tight frame denoising method based on the data-driven tight frame construction scheme. The experiments showed that our adaptive tight frame denoising technique significantly outperformed standard wavelet thresholding approaches on images of rich textures. Also, it is much faster than some over-complete dictionary based approaches, e.g. the K-SVD method [11], with comparable performance. The rest of the paper is organized as follows. In Section 2, we first give a brief introduction to the preliminaries of wavelet tight frame. Then, in Section 3, we introduce the proposed minimization model and the corresponding numerical method. Section 4 is devoted to the experimental evaluation of the proposed method and discussions.

2. Preliminaries and previous work

2.1. Wavelet and tight frame

In this section, we give a brief introduction to tight frame, wavelet tight frame in a Hilbert space $\mathcal{H}$ and their constructions. Interested readers are referred to [3, 14, 15] for more details. A sequence $\{x_n\} \subset \mathcal{H}$ is a *tight frame* for H if

$$\|x\|^2 = \sum_n |\langle x, x_n\rangle|^2, \quad \text{for any } x \in \mathcal{H}.$$

There are two associated operators. One is the analysis operator W defined by

$$W : x \in \mathcal{H} \longrightarrow \{\langle x, x_n\rangle\} \in \ell_2(\mathbb{N})$$

and the other is its adjoint operator $W^\top$ called the synthesis operator:

$$W^\top : \{a_n\} \in \ell(\mathbb{N}) \longrightarrow \sum_n a_n x_n \in \mathcal{H}.$$

Then, a sequence $\{x_n\} \subset \mathcal{H}$ is a tight frame if and only if $W^\top W = I$, where $I : \mathcal{H} \longrightarrow \mathcal{H}$ is the identical operator. In other words, given a tight frame $\{x_n\}$, we have the following canonical expansion:

$$x = \sum_n \langle x, x_n \rangle x_n, \quad \text{for any } x \in \mathcal{H}.$$

The sequence $\{\langle x, x_n \rangle\}$ is called the *canonical* tight frame coefficient sequence. Thus, tight frames are often viewed as generalizations of orthonormal bases. In fact, a tight frame $\{x_n\}$ is an orthonormal basis for $\mathcal{H}$ if and only if $\|x_n\| = 1$ for all x_n.

One widely used class of tight frames in signal processing is the *wavelet tight frame*. The wavelet tight frame for $L_2(\mathbb{R})$ starts with a finite set of generators $\Psi := \{\psi^2, \psi^3, \ldots, \psi^m\} \in L_2(\mathbb{R})$. The shifts and dilations of the generators defines an affine system:

$$X(\Psi) = \{\psi^\ell_{j,k}, 2 \le \ell \le m, j \in \mathbb{Z}, k \in \mathbb{Z}\},$$

where $\psi^\ell_{j,k} = M^{j/2}\psi^\ell(M^j \cdot -k)$ for some positive integer M. The affine system $X(\Psi) \subset L_2(\mathbb{R})$ is called a *wavelet tight frame* of $L_2(\mathbb{R})$ if it is a tight frame satisfying

$$f = \sum_{x \in X(\Psi)} \langle f, x \rangle x, \quad \text{for any } x \in L_2(\mathbb{R}). \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the inner product of $L_2(\mathbb{R})$. One construction scheme of wavelet tight frame systems is using multi-resolution analysis (MRA). The construction of MRA-based wavelet tight frames usually starts with a compactly supported refinable function ϕ (often referred as a scaling function) with a refinement mask a_1 satisfying

$$\widehat{\phi}(M\cdot) = \widehat{a_1}\widehat{\phi},$$

where $\widehat{\phi}$ is the Fourier transform of ϕ, and $\widehat{a_1}$ is a 2π-periodic trigonometric polynomial defined as $\widehat{a_1}(\omega) := \sum_{k\in\mathbb{Z}} a_1(k)e^{-ik\omega}$ and $\widehat{a_1}(0) = 1$. After obtaining a compactly supported refinable function ϕ, the next step is to find an appropriate set of framelets $\Psi = \{\psi_2, \ldots, \psi_m\}$ defined in the Fourier domain by

$$\widehat{\psi_i}(M\cdot) = \widehat{a_i}\widehat{\phi}, \quad i = 2, 3, \ldots, m$$

such that $X(\Psi)$ forms a wavelet tight frame, where the associated masks $\{a_i\}_{i=2}^m$ are 2π-periodic trigonometric polynomials:

$$\widehat{a_i}(\omega) = \sum_{k\in\mathbb{Z}} a_i(k)e^{-i\pi k\omega}.$$

The construction of MRA-based wavelet tight frames is based on the Unitary Extension Principle (UEP) [2], which says that $X(\Psi)$ forms a tight frame provided that $\phi \in L_2(\mathbb{R})$ and

$$\sum_{i=1}^{m} \widehat{a_i}(\omega)\overline{\widehat{a_i}(\omega + 2\pi\gamma)} = \delta_\gamma, \quad \text{for } \gamma \in \Omega_M \cap [0,1), \tag{2}$$

where $\Omega_M := (M)^{-1}\mathbb{Z}$. A class of MRA-based wavelet tight frames are constructed in [2, 3] using the UEP. For example, the linear B-spline framelet used in many image restoration tasks (see e.g. [16, 17, 18]) has the following three masks:

$$a_1 = \frac{1}{4}(1,2,1)^\top; \quad a_2 = \frac{\sqrt{2}}{4}(1,0,-1)^\top; \quad a_3 = \frac{1}{4}(-1,2,-1)^\top. \tag{3}$$

Once the wavelet tight frame for $L_2(\mathbb{R})$ is obtained, one way to construct higher-dimensional wavelet tight frame for $L_2(\mathbb{R}^2)$ is via the tensor product.

The masks of the tight frame constructed for $L_2(\mathbb{R}^2)$ can be used to generate the tight frame for the space of square summable sequences, denoted by $\ell_2(\mathbb{Z}^2)$. Given some positive integer M, let $\downarrow_M$ denote the down-sampling operator and $\uparrow_M$ the up-sampling operator such that

$$\begin{aligned} &[v\downarrow_M](n) := v(Mn), \quad n \in \mathbb{Z}^2; \\ &[v\uparrow_M](Mn) := v(n), \quad n \in \mathbb{Z}^2, \quad \text{and all other elements of } v\uparrow_M \text{ are zero.} \end{aligned}$$

Then there are two essential operators associated with wavelet tight frames. One is the *decomposition* operator $T_a : \ell_2(\mathbb{Z}^2) \to \ell_2(\mathbb{Z}^2)$ defined by

$$[T_a v](n) := M[(a(-\cdot) * v)\downarrow_M](n) = M\sum_{k\in\mathbb{Z}^2} v(k)a(k-Mn), \quad \forall\, v \in \ell_2(\mathbb{Z}^2), \tag{4}$$

and the other is the *reconstruction* operator $R_a : \ell_2(\mathbb{Z}^2) \to \ell_2(\mathbb{Z}^2)$ defined by

$$[R_a v](n) := M[(a * (v\uparrow_M)](n) = M\sum_{k\in\mathbb{Z}^2} v(k)\overline{a(n-Mk)}, \quad \forall\, v \in \ell_2(\mathbb{Z}^2), \tag{5}$$

where '$*$' denote the discrete convolution operator. Let $\{a_i \in \ell_2(\mathbb{Z}^2)\}_{i=1}^m$ denote the masks (filters) associated with a wavelet tight frame system. If considering a single-level wavelet tight frame decomposition and reconstruction, then

Theorem 1. ([19]) *Let $a_j, j = 1, \ldots, m$ be m be m real-valued finitely supported filters. Then the following three conditions are equivalent:*

(i) $\sum_{i=1}^m \widehat{a_i}(\omega)\overline{\widehat{a_i}(\omega + 2\pi\gamma)} = \delta_\gamma$, *for any* $\gamma \in M^{-1}\mathbb{Z}^2 \cap [0,1)^2$;

(ii) $\sum_{i=1}^m R_{a_i}T_{a_i}v = v$, *for any* $v \in \ell_2(\mathbb{Z}^2)$;

(iii) $\sum_{i=1}^m \sum_{n\in\mathbb{Z}^2} a_i(k+Mn+\ell)a_i(Mn+\ell) = M^{-1}\delta_k$, *for any* $k, \ell \in \mathbb{Z}^2$.

If the wavelet tight frame system is a shift-invariant system by using $M = I$, then there is no downsampling involved in (4) and (5). Let $W^\top$ be the set formed by all sequences $v_{a_i,k}$ given by

$$[v_{a_i,k}](n) = a_i(n-k), \quad n \in \mathbb{Z}^2, \tag{6}$$

for $k \in \mathbb{Z}^2$. Suppose that $\{a_i\}_{i=1}^m$ is the set of all filters associated with a tight frame in $L_2(\mathbb{R}^2)$ satisfying the UEP condition. Then the equivalence between (i) and (ii) in Theorem 1 gives

$$W^\top W = I,$$

In other words, the elements of $W^\top$ form a tight frame of $\ell_2(\mathbb{Z}^2)$. Also, by the equivalence between (i) and (iii) in Theorem 1, the UEP condition (2) for the shift-invariant wavelet tight frame can be expressed as the following,

$$\sum_{i=1}^{m} \sum_{n \in \mathbb{Z}^2} a_i(k+n) a_i(n) = \delta_k, \quad \text{for any } k \in \mathbb{Z}^2. \tag{7}$$

The analysis operator W and the synthesis operator $W^\top$ associated with a tight frame for sequences are also applicable to images, finite sequences in $\ell_2(\mathbb{Z}^2)$. Let $f \in \mathbb{R}^N$ denote the vector form of the image $\mathbf{f}$ by concatenating all columns of $\mathbf{f}$ into a single column vector. For a given finitely supported 2D filter a, let the N-by-N matrix $\mathcal{S}_a$ be the matrix that represents the decomposition operator R_a under Neumann boundary condition (see [17, 14] for more details). Then, the synthesis operator $W^\top \in \mathbb{R}^{N \times P}$ for $\mathbb{R}^N$ w.r.t. the wavelet filters $H = \{a_k\}_{k=1}^m$ is defined as

$$W^\top = [\mathcal{S}_{a_1}, \mathcal{S}_{a_2}, \ldots, \mathcal{S}_{a_m}]. \tag{8}$$

The columns of $W^\top$ form a wavelet tight frame for $\mathbb{R}^N$ such that $W^\top W = I_N$. Interesting readers are referred to [14] for more details.

2.2. Previous work on adaptively learning over-complete dictionaries

In the past, there have been several approaches proposed to learn an over-complete dictionary from a signal/image such that the sparsity of the target signal/image under the learned over-complete dictionary is optimized, e.g., [20, 9, 10, 11, 12, 13]. Earlier works are based on probabilistic reasoning. For example, the maximum likelihood method is used in [20, 9] and the maximum A-posteriori probability approach is used in [10] to construct over-complete dictionaries. In recent years, there have been steady progresses on development of the deterministic approaches of learning an over-complete dictionary from the image (e.g. [11, 12, 13]). The representative work along this direction is the so-called K-SVD method [11], which presents a minimization model to learn a dictionary from the noisy image and use the learned dictionary to denoise images. Since our work also belongs to this category, we only give a brief review on the K-SVD method.

Let f, g denote the vector forms of the images $\mathbf{f}$ and $\mathbf{g}$ respectively. Let $D \in \mathbb{R}^{m,p}$ denote the dictionary whose column vectors denotes the dictionary atoms. Partitioning the image into ℓ image patches of m pixels with overlaps. Then let P_j denotes the projection operator that maps the image to its j-th image for $j = 1, 2, \ldots, \ell$. Then the K-SVD method is to denoise image by solving the following minimization:

$$\min_{g,c,D} \frac{1}{2}\|f - g\|_2^2 + \sum_j \mu_j \|c_j\|_0 + \lambda \sum_j \|Dc_j - P_j g\|_2^2, \tag{9}$$

where $f, g \in \mathbb{R}^n$ are the noisy image and the noise-free image to be estimated, $c = (c_1, c_2, \ldots, c_\ell) \in \mathbb{R}^{p,\ell}$ is the matrix whose j-th column vector c_j denotes the expansion coefficient vector of j-th image patch over the dictionary D. The K-SVD method showed better performance than the standard wavelet thresholding method in the application of image denoising. Such an improvement comes from the fact that the repeating texture elements are likely to be captured as atoms in the system learned by the method.

Despite the impressive performance of the K-SVD method in image denoising, how to efficiently solve the minimization (9) with satisfactory stability remains a challenging task, as the minimization (9) is a very challenging ill-posed minimization problem. An alternating iterative method is implemented in [11] to alternatively update the estimations of c, g and D during each iteration. Under such an scheme, there are two challenging sub-problems to solve during each iteration. One is how to estimate the dictionary D given the current estimations on c and g, which is a severely under-constrained problem owing to the redundancy of D ($p \gg m$). A heuristic method is proposed in [11] to update the atoms of the dictionary one by one in a greedy manner, which lacks the rigorous treatment on the stability and optimality. Also, how to find a sparse coefficient c given an over-complete dictionary D could be a computationally expensive process. The orthogonal matching pursuit is used in [11] to find a sparse coefficient vector c during each iteration, which is quite slow and accounts for most of computational amount of the K-SVD method.

3. Minimization model and numerical method

In this section, we present a new minimization model for constructing an adaptive discrete tight frame for the given image. An efficient numerical solver for solving the proposed minimization model is also provided.

3.1. Basic idea

Same as wavelet tight frame, the discrete tight frame constructed in our approach is also generated by the shifts of a few generators. Moreover, it is known that a shift-invariant system is more effective on reducing the artifacts of the results than a shift-variant system in many image restoration tasks (see e.g. [21, 14]). Thus, the tight frame constructed in our approach is a shift-invariant system. In other words, the tight frame constructed in our approach

is in the form of (8). Suppose that the image is of size $L \times M$ and let $N = LM$. Let $\mathcal{S}_a : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the block-wise Toeplitz matrix that represents the convolution operator with a finitely supported 2D filter a under Neumann boundary condition. Then the synthesis operator $W^\top \in \mathbb{R}^{N \times Nm}$ is defined by

$$W^\top = [\mathcal{S}_{a_1}, \mathcal{S}_{a_2}, \ldots, \mathcal{S}_{a_m}], \tag{10}$$

where $\{a_i\}_{i=1}^m$ are the filters associated with a tight frame and the columns of $W^\top$ form a tight frame for $\mathbb{R}^N$. Its transpose, the analysis operator $W \in \mathbb{R}^{Nm \times N}$ is then as follows,

$$W = [\mathcal{S}_{a_1(-\cdot)}^\top, \mathcal{S}_{a_2(-\cdot)}^\top, \ldots, \mathcal{S}_{a_m(-\cdot)}^\top]^\top. \tag{11}$$

Let $g \in \mathbb{R}^N$ denote an input image and let W denote the analysis operator of the tight frame defined by (11). We propose to construct a tight frame $W^\top(a_1, \ldots, a_m)$ by solving the following minimization:

$$\min_{v, \{a_i\}_{i=1}^m} \|v - W(a_1, a_2, \ldots, a_m)g\|_2^2 + \lambda^2 \|v\|_0, \quad \text{subject to} \quad W^\top W = I. \tag{12}$$

There are two unknowns: one is the coefficient vector v which sparsely approximates the canonical tight frame coefficient Wg, and the other is the set of filers $\{a_i\}_{i=1}^m$ that generates a tight frame as (10). We take an iterative scheme to alternatively update the estimation of the coefficient vector v and the estimation of $\{a_i\}_{i=1}^m$. More specifically, let $\{a_i^{(0)}\}_{i=1}^m$ be the set of the initial filters to start with, e.g. the linear spline framelet filters. Then for $k = 0, 1, \ldots, K-1$,

1. Given the frame filters $\{a_i^{(k)}\}_{i=1}^m$, define the sparse frame coefficient vector $v^{(k)}$ by
$$v^{(k)} := \text{argmin}_v \|v - W^{(k)}g\|_2^2 + \lambda^2 \|v\|_0, \tag{13}$$
where $W^{(k)}$ is the analysis operator derived from $\{a_i^{(k)}\}_{i=1}^m$ as (11).

2. Given the sparse frame coefficient vector $v^{(k)}$, update the frame filters $\{a_i^{(k+1)}\}_{i=1}^m$ by
$$\{a_i^{(k+1)}\}_{i=1}^m := \text{argmin}_{\{a_i\}_{i=1}^m} \|v^{(k)} - W(a_1, a_2, \ldots, a_m)g\|_2^2 \tag{14}$$
subject to $W^\top W = I_N$, where $W^\top$ is the tight frame defined by $\{a_i\}_{i=1}^m$ as (10).

After $K+1$ iterations, the tight frame adaptive to the image g is defined as $W^\top(a_1^{(K)}, a_2^{(K)}, \ldots, a_m^{(K)})$. In the next, we will give a more detailed discussion on the model (12) and present an efficient numerical solver.

3.2. Minimization (13) *and its implications*

It is known that the minimization (13) has a unique solution $\bar{v}$ by applying a hard thresholding operator on the canonical frame coefficient vector Wg:

$$\bar{v} := T_\lambda(Wg), \tag{15}$$

where $T_\lambda : \mathbb{R}^{Nm} \to \mathbb{R}^{Nm}$ denote the *hard thresholding* operator defined by

$$[T_\lambda v](n) = \begin{cases} v(n) & \text{if } |v(n)| > \lambda; \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Proposition 2. *Let $\bar{v}$ denote the solution to the minimization* (13)*. Then, for all v with $\|v\|_0 \le \|\bar{v}\|_0$,*

$$\|v - Wg\|_2 \ge \|\bar{v} - Wg\|_2.$$

Proof. For any $v \in \mathbb{R}^{Nm}$, by the definition of $\bar{v}$, we have

$$\|v - Wg\|_2^2 + \lambda^2\|v\|_0 \ge \|\bar{v} - Wg\|_2^2 + \lambda^2\|\bar{v}\|_0.$$

Thus

$$\|v - Wg\|_2^2 - \|\bar{v} - Wg\|_2^2 = \lambda^2(\|\bar{v}\|_0 - \|v\|_0) \ge 0.$$

The proof is done.

Proposition 2 states that $\bar{v}$ is the best approximation to the canonical frame coefficient vector Wg among all sparse vectors whose cardinality no greater than the cardinality of $\bar{v}$. Using the coefficient $\bar{v}$, we can reconstruct a signal $\bar{g}$ defined by

$$\bar{g} = W^\top \bar{v} = W^\top(T_\lambda(Wg)).$$

When $W^\top$ is a redundant tight frame, the approximation error of the reconstruction $\bar{g}$ to g is bounded by $\|\bar{v} - Wg\|_2$. However, $\bar{g}$ is not necessarily the best approximation to the signal g.

Proposition 3. *Let W denote the tight frame satisfying $W^\top W = I$. Then*

$$\|g - W^\top v\|_2^2 + \|(I - WW^\top)v\|_2^2 = \|v - Wg\|_2^2.$$

Proof. By the fact that $W^\top W = I$, we have

$$\begin{aligned} &\|g - W^\top v\|_2^2 + \|(I - WW^\top)v\|_2^2 \\ &= g^\top g - 2g^\top W^\top v + v^\top WW^\top v + v^\top (I - WW^\top)(I - WW^\top)v \\ &= g^\top g - 2g^\top W^\top v + v^\top WW^\top v + v^\top (I - WW^\top)v \\ &= g^\top g - 2g^\top W^\top v + v^\top v \\ &= g^\top W^\top Wg - 2g^\top W^\top v + v^\top v \\ &= \|v - Wg\|_2^2. \end{aligned} \tag{17}$$

This completes the proof.

Proposition 3 essentially says that reconstruction $\bar{g}$ from the coefficients $\bar{v}$ obtained via (13) indeed minimizes the following objective function:

$$\|g - W^\top v\|_2^2 + \|(I - WW^\top)v\|_2^2 + \lambda^2\|v\|_0. \tag{18}$$

The minimization (18) is actually the so-called *balanced* approach for sparsity-based regularizations (see more details in [14]). In the next, we give a brief discussion on such a balance approach. Given a signal f and a tight frame system $W^\top$, there exist several regularizations to find a sparse approximation to f with different outcomes. Suppose we are using ℓ_1 norm as a convex replacement of ℓ_0 norm to prompt sparsity. Most existing regularization methods of finding a sparse approximation $\bar{g}$ to the signal g is done via solving the following minimization:

$$\bar{g} := W^\top \bar{v}; \quad \bar{v} := \text{argmin}_v \|g - W^\top v\|_2^2 + \tau\|(I - WW^\top)v\|_2^2 + 2\lambda\|v\|_1, \tag{19}$$

Based on different values of τ, the minimization (19) can be classified into three categories. When $\tau = 0$, the model (19) is called the *synthesis* based approach (e.g. [22, 18]). When $\tau = \infty$, the minimization (19) can be rewritten as

$$\bar{g} := \text{argmin}_f \|g - f\|_2^2 + 2\lambda\|Wg\|_1 \tag{20}$$

as the coefficient $v \in \text{range}(W)$. The above approach is called the *analysis* based approach (e.g. [23, 24]). When $0 < \tau < \infty$, the model (19) is called a *balanced* approach (e.g. [25, 16, 17]).

The main difference among three methods lies in how much the second term $\|(I - WW^\top)v\|_2^2$ contributes to the objective function. By rewriting $\|(I - WW^\top)v\|_2^2$ as $\|v - W(W^\top v)\|_2^2$, we see that it measures the distance between the coefficient vector v and the canonical coefficients of its corresponding reconstructed signal $W^\top v$. Since the magnitude of the canonical coefficients reflects the regularity of the reconstructed signal under some mild conditions on the tight frame $W^\top$ (see [26] fore more details), this distance is closely related to the regularity of the reconstructed signal. The smaller is $\|(I - WW^\top)v\|_2^2$, the more accurately the corresponding coefficient v reflects the regularity of the underlying signal. Thus, the synthesis based approach emphasizes the sparsity of the coefficient vector $\bar{v}$, but the decay of $\bar{v}$ does not reflect the regularity of the resulting reconstruction $W^\top \bar{v}$. In contrast, the analysis based approach emphasizes the sparsity of the canonical coefficient vector $W\bar{g}$ which leads to a more regular approximation. Indeed, it is shown in [27] that, by choosing parameters properly, the analysis based approach can be seen as sophisticated discretization of minimizations involving the total variation penalties or their generalizations. It is noted that the coefficient vector $\bar{v}$ obtained from the synthesis based approach will be much more sparse than the canonical coefficient vector $W\bar{g}$ obtained by the analysis-based approach. The balanced approach yields the result which balances the sparsity of the coefficient vector and the regularity of the reconstructed signal.

Each approach has its disadvantages and advantages. The choice of the approach depends on the nature of the targeted application. It is empirically

observed that for image restoration, the reconstructed image obtained by the synthesis-based approach tends to have some visually unpleasant artifacts. In contrast, the results from the analysis-based approach and the balanced approach usually have less artifacts as they ensure certain regularities along image edges. For our purpose, we take the balanced approach and the reason is two-fold. First, our goal is to construct the tight frame that works better for image restoration than the existing ones. Thus, we are not seeking for the tight frame that maximizes the sparsity of the coefficients. Second, the minimization model (20) resulting from the analysis-based approach requires an iterative solver, which is too expensive for our purpose as the minimization (20) will be called for many times in our approach. Therefore, we take an balanced approach and the result $W^\top(T_\lambda(Wg)$ seeks the balance between the regularity of the result and the sparsity of its associated tight frame coefficients.

3.3. The modification of Model (14) *and its numerical solver*

The minimization model (14) is a constrained minimization with the quadratic constraints $W^\top W = I$. Since the tight frame we are interested is a shift-invariant system, the UEP condition (2) for ensuring $W^\top W = I$ is then simplified to the constraints (7):

$$\sum_{i=1}^{m}\sum_{n\in\mathbb{Z}^2} a_i(k+n)a_i(n) = \delta_k, \quad \text{for any } k\in\mathbb{Z}^2. \tag{21}$$

The minimization (14) with the above quadratic constraints requires solving a rather complex non-convex minimization problem. Thus, we propose to construct the tight frame with certain special structure which will greatly simplify the quadratic constraints (7) such that the minimization (14) has an explicit analytic solution.

Proposition 4. *Let $a_i, i = 1, 2, \ldots, r^2$ be r^2 real-valued filters with support on $\mathbb{Z}^2 \cap [1, r]^2$ for some positive integer r. Then the filters $\{a_i\}_{i=1}^{r^2}$ satisfy the following condition:*

$$\sum_{i=1}^{r^2}\sum_{n\in\mathbb{Z}^2} a_i(k+n)a_i(n) = \delta_k, \quad k\in\mathbb{Z}^2,$$

as long as they satisfy the following orthogonal constraints:

$$\langle a_i, a_j\rangle = \sum_{k\in[1,r]^2\cap\mathbb{Z}^2} a_i(k)a_j(k) = \frac{1}{r^2}\delta_{i-j}, \quad 1\le i,j\le r^2.$$

Proof. For each a_i, let $\vec{a}_i$ denote its column vector form by concatenating all its columns. Define the matrix $A\in\mathbb{R}^{r^2\times r^2}$ by

$$A = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{r^2}]. \tag{22}$$

Then

$$(A^\top A)(i,j) = \vec{a}_i^\top \vec{a}_j = \frac{1}{r^2}\delta_{i-j}, \quad 1 \le i,j \le r^2.$$

Thus we have $A^\top A = \frac{1}{r^2} I$ which implies $AA^\top = \frac{1}{r^2} I$. Notice that

$$(AA^\top)(k,\ell) = \sum_{i=1}^{r^2} \vec{a}_i(k)\vec{a}_i(\ell) = \frac{1}{r^2}\delta_{k-\ell}, \quad 1 \le k,\ell \le r^2,$$

which gives $\sum_{i=1}^{r^2} a_i(m)a_i(n) = \frac{1}{r^2}\delta_{m-n}, m,n \in \mathbb{Z}^2$. Then we have

$$\begin{aligned}
\sum_{i=1}^{r^2}\sum_{n\in\mathbb{Z}^2} a_i(k+n)a_i(n) &= \sum_{n\in\mathbb{Z}^2}\sum_{i=1}^{r^2} a_i(k+n)a_i(n) \\
&= \begin{cases} \frac{r^2}{r^2} = 1, & k = \mathbf{0}; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

The proof is done.

By considering the frame filters $\{a_i\}_{i=1}^{r^2}$ with a special structure proposed in Proposition 4, we greatly simplified the quadratic constraints on $\{a_i\}_{i=1}^{r^2}$ to ensure that $W^\top W = I$. The minimization (14) is now simplified to

$$\{a_i^{(k+1)}\}_{i=1}^{r^2} := \text{argmin}_{\{a_i\}_{i=1}^{r^2}} \|v^{(k)} - W(a_1, a_2, \ldots, a_{r^2})g\|_2 \tag{23}$$

subject to $\langle a_i, a_j\rangle = \frac{1}{r^2}\delta_{i-j}$, $1 \le i,j \le r^2$, where W is the analysis operator defined by $\{a_i\}_{i=1}^{r^2}$ in the form of (11).

In the next, we derive the explicit analytic solution to the constrained minimization problem (23). Sequentially partitioning the coefficient vector $v^{(k)}$ into r^2 vectors, denoted by $v^{(k),i} \in \mathbb{R}^{N\times 1}, i = 1, 2 \ldots, r^2$. Then we can rewrite the objective function of (23) as follows,

$$\begin{aligned}
\|v^{(k)} - Wg\|_2^2 &= \sum_{i=1}^{r^2} \|v^{(k),i} - \mathcal{S}_{a_i(-\cdot)}g\|_2^2 \\
&= \sum_{i=1}^{r^2}\sum_{n=1}^{N} \|v^{(k),i}(n) - [\mathcal{S}_{a_i(-\cdot)}g](n)\|_2^2 \\
&= \sum_{n=1}^{N}\sum_{i=1}^{r^2} \|v^{(k),i}(n) - [\mathcal{S}_{a_i(-\cdot)}g](n)\|_2^2.
\end{aligned}$$

Let $\vec{a}_i$ denote the vector form of a_i by concatenating all its columns. Since the convolution is commutative, for $i = 1, 2, \ldots, r^2$, we have then

$$[\mathcal{S}_{a_i(-\cdot)}g](n) = [\mathcal{S}_{g(-\cdot)}\vec{a}_i](n) = \vec{g}_n^\top \vec{a}_i = \vec{a}_i^\top \vec{g}_n, \quad 1 \le n \le N,$$

where $\vec{g}_n$ denote the transpose of the n-th row of $\mathcal{S}_{g(-\cdot)}$. Let

$$\vec{v}_n = \left(v^{(k),1}(n), v^{(k),2}(n), \ldots, v^{(k),r^2}(n)\right)^\top, 1 \leq n \leq N,$$

and define

$$\begin{cases} V = (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) \in \mathbb{R}^{r^2 \times N} \\ G = (\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_N) \in \mathbb{R}^{r^2 \times N} \\ A = (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{r^2}) \in \mathbb{R}^{r^2 \times r^2}. \end{cases} \tag{24}$$

We have then

$$\begin{aligned} \|v^{(k)} - W^\top g\|_2^2 &= \sum_{n=1}^{N} \|\vec{v}_n - A^\top \vec{g}_n\|_2^2 \\ &= \sum_{n=1}^{N} \vec{v}_n^\top \vec{v}_n + \vec{g}_n^\top A A^\top \vec{g}_n - 2\vec{v}_n^\top A^\top \vec{g}_n \\ &= \sum_{n=1}^{N} \vec{v}_n^\top \vec{v}_n + \frac{1}{r^2} \vec{g}_n^\top \vec{g}_n - 2(A\vec{v}_n)^\top \vec{g}_n \\ &= \mathrm{Tr}(V^\top V) + \frac{1}{r^2}\mathrm{Tr}(G^\top G) - 2\mathrm{Tr}(AVG^\top) \end{aligned}$$

where $\mathrm{Tr}(\cdot)$ denote the trace of the matrix. Since the first two terms are constants, the minimization (23) can be rewritten as follows,

$$\max_A \mathrm{Tr}(AVG^\top) \quad \text{s.t.} \quad A^\top A = \frac{1}{r^2} I_{r^2}. \tag{25}$$

The following theorem gives an explicit solution to the above minimization (25).

Theorem 5. ([28]) *Let B and C be $m \times r$ matrices and B has rank r. Consider the constrained maximization problem:*

$$B_* = \mathrm{argmax}_B \mathrm{Tr}(B^\top C), \quad s.t.\ B^\top B = I_r,$$

Suppose that the single value decomposition (SVD) of C is $C = UDX^\top$, then $B_* = UX^\top$.

By Theorem 5, we obtained the solution of (25):

$$A_* = \frac{1}{r}(UX^\top)^\top = \frac{1}{r} XU^\top, \tag{26}$$

where U and X are the SVD decomposition of $VG^\top$ such that

$$VG^\top = UDX^\top.$$

In other words, the vector form of the filters $a_i^{(k+1)}$ defined by the minimizer of (23) is exactly the i-th column vector of the matrix A_* given by (26).

In summary, the ultimate minimization model we proposed for constructing a tight frame adaptive to the input image is as follows,

$$\{a_i^*\}_{i=1}^{r^2} := \text{argmin}_{v,\{a_i\}_{i=1}^{r^2}} \|v - W(a_1, \ldots, a_{r^2})g\|_2^2 + \lambda^2 \|v\|_0 \tag{27}$$

subject to $\langle a_i, a_j \rangle = \frac{1}{r^2}\delta_{i-j}$, $1 \le i, j \le r^2$, where W is the analysis operator defined by $W = [\mathcal{S}_{a_1(-\cdot)}^\top, \mathcal{S}_{a_2(-\cdot)}^\top, \ldots, \mathcal{S}_{a_r(-\cdot)}^\top]^\top$. The complete description of the numerical solver for solving (27) is given in Algorithm 1. Notice that there are two steps during each iteration and each involves solving one minimization problem. The first is simply done by applying a hard thresholding operator on tight frame coefficients $W(a_1^{(k)}, \ldots, a_{r^2}^{(k)})g$, and the second can be obtained by the single value decomposition of the matrix $VG^\top$. Thus, the computation cost for each iteration is very low.

Algorithm 1 Construction of the discrete tight frame adaptive to the image

Input: the image g (clean or noisy)

Output: a shift-invariant discrete tight frame $W^\top$ defined by filters $\{a_i^{(K)}\}_{i=1}^{r^2}$

Main procedure:

(I) Initialize tight frame filters $\{a_i^{(0)}\}_{i=1}^{r^2}$ of size at most $r \times r$ by some existing tight frame system.

(II) For $k = 0, 1, 2, \ldots, K-1$ do

(1) define $W^{(k)}$ from $\{a_i^{(k)}\}_{i=1}^{r^2}$ by (11);

(2) set $v^{(k)} = T_\lambda(W^{(k)}g)$, where T_λ is the hard thresholding operator defined in (16);

(3) assembling the matrix V, G by (24);

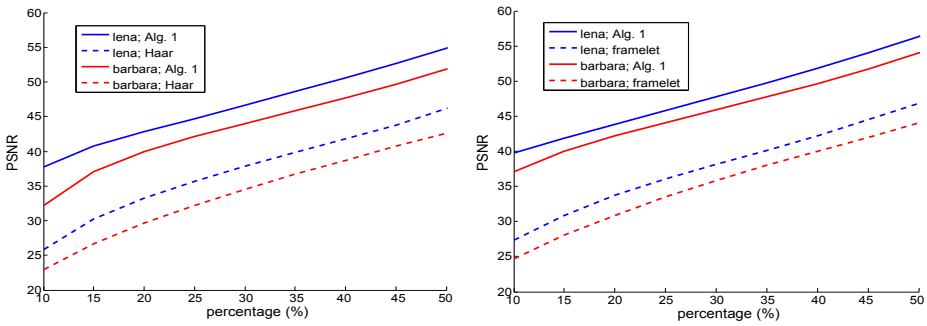
(4) run the SVD decomposition on $VG^\top$ s.t. $VG^\top = UDX^\top$;

(5) set $a_i^{(k+1)}$ to the i-th column vector of the matrix $A^{(k+1)} = \frac{1}{r}XU^\top$ for $i = 1, \ldots, r^2$.

(III) Output $\{a_i^{(K)}\}_{i=1}^{r^2}$.

3.4. Experiments on some real images

To illustrate the behavior of the tight frame from Algorithm 1, the proposed algorithm is tested on some real images which contains both cartoon-type regions and texture regions. As we discussed in Section 3.2, neither Algorithm 1 is seeking for the tight frame which maximizes the sparsity of the canonical tight frame coefficients, nor it is seeking for the tight frame in which the signal can be approximated by a most sparse coefficient vector. Instead, Algorithm 1 seeks for the tight frame system whose resulting approximation balances the sparsity



(a) (b)

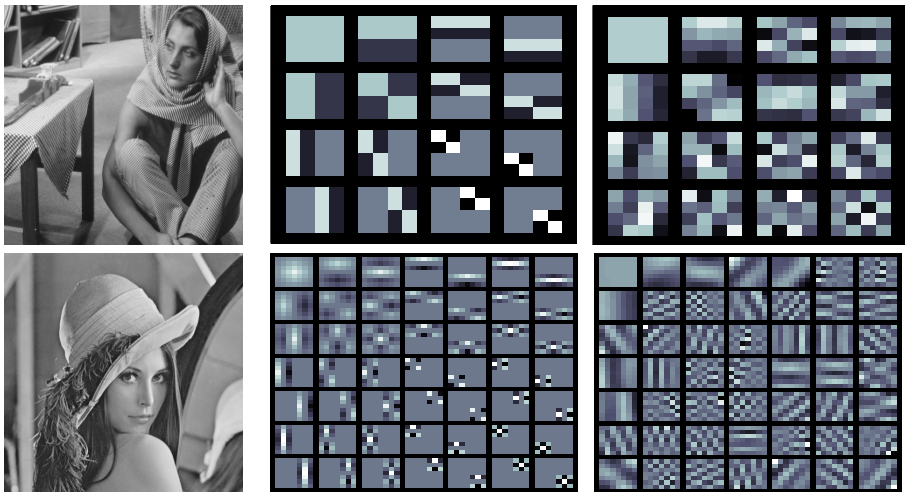
Figure 1: The comparison of the PSNR values of the reconstructions using only a portion of the tight frame coefficients under different tight frame systems. The x-axis denotes the percentage of the tight frame coefficients used in the reconstruction and the y-axis denotes the PSNR value of the reconstruction. (a) The comparison between the 5-level shift-invariant Haar wavelet system and the tight frame systems obtained by Algorithm 1 taking the input as the Haar filters. Both system have 16 filters in total. (b) The comparison between the 6-level shift-invariant linear spline framelet system and the tight frame systems obtained by Algorithm 1 taking the input as the linear framelet filters. Both system have 49 filters in total.

of tight frame coefficients and the regularity of the reconstruction. Nevertheless, the tight frame reconstructed from Algorithm 1 is still much more effective on sparsifying the canonical tight frame coefficients than the existing non-adaptive ones. In the experiments, we measure the effectiveness of such a sparsification as follows. Given a tight frame W and an image g, let $\Gamma_{\alpha\%}$ denote the hard thresholding operator which keeps $\alpha\%$ of largest frame coefficients in absolute value and set all other coefficients zero. Then we calculate the PSNR value of $W^\top(\Gamma_{\alpha\%}(Wg))$ to measure the quality of the reconstructed image by only using $\alpha\%$ of frame coefficients. Using the same percentage of canonical frame coefficients, the larger is the PSNR value of the reconstruction, the more effective the tight frame sparsifies the canonical coefficients of the input image. For an image $\mathbf{x}$, the peak signal to noise ratio (PSNR) of its estimate $\widehat{\mathbf{x}}$ is defined as

$$\mathrm{PSNR}(\widehat{\mathbf{x}},\mathbf{x}) = 10\log_{10}\frac{255^2}{\frac{1}{LM}\sum_{i=1}^{L}\sum_{j=1}^{M}(\widehat{\mathbf{x}}(i,j)-\mathbf{x}(i,j))^2}$$

where L and M are the dimensions of the image $\mathbf{x}$, and $\mathbf{x}(i,j), \widehat{\mathbf{x}}(i,j)$ are the pixel value of the input and the estimate image at the pixel location (i,j).

In the experiments, Algorithm 1 is applied on two images "Barbara" and "Lena" shown in Fig. 2 (a) with two different initializations. The filters of one initialization is the tensor Haar wavelet filters with totally 16 filters and the filters of another initialization is the tensor linear spline framelet [3] with totally 49 filters. Through all experiments, The maximum iteration number K of Algorithm 1 is set to 25. In Figure 1, the PSNR values of the reconstruction using different percentages of tight frame coefficients are plotted with



(a) (b) (c)

Figure 2: The illustration of the data-driven tight frame filters constructed by Algorithm 1. (a) Two tested images "Barbara" and "Lena"; (b): The filters associated with the multi-level Haar wavelets (on the top) and the filters associated with the multi-level linear spline framelet (on the bottom); (c) the corresponding adaptive tight frame filters constructed by Algorithm 1 using the filters in (b) as the input. One small block of the images shown in (b) and (c) represents one filter.

respect to different images and different tight frame systems. Clearly, the tight frames constructed by Algorithm 1 are more effective on sparsifying the canonical tight frame coefficients of the input images than the Haar wavelets and linear framelets, as the PSNR values of the reconstructed images are significantly higher. Such an effectiveness comes from the fact that the tight frame constructed by Algorithm 1 can efficiently capture the repeating complex texture patterns in two input images while the two existing frame systems can not. See Fig. 2 for an illustration of the filters of the tight frames constructed by Algorithm 1 with respect to two different initializations and two different images. It is seen that the filters constructed by Algorithm 1 tends to fit the repeating patterns in the images.

4. Adaptive tight frame image denoising method

Algorithm 1 can be easily extended to an adaptive tight frame denoising method as follows. The proposed construction scheme of adaptive tight frame is first applied to a noisy input and then the obtained tight frame is used for noise removal. Let $f = g+n$ denote some noisy observation of g where n denotes the i.i.d. Gaussian white noise. Then the first two steps in (II) of Algorithm 1 can be viewed as a thresholding denoising method under a tight frame. The

intermediate denoising result during each iteration is then

$$g^{(k)} = {W^{(k)}}^{\top}(T_{\lambda}(W^{(k)}f))$$

with $W^{(k)} = W(a_1^{(k)}, \ldots, a_{r^2}^{(k)})$, where λ is the threshold whose value depends on both the noise level and the desired sparsity degree of the image. After Algorithm 1 is terminated. The resulting tight frame $W^{(K)} = W(a_1^{(K)}, \ldots, a_{r^2}^{(K)})$ then can be used to denoise the image.

$$g^{(K)} = {W^{(K)}}^{\top}(T_{\widetilde{\lambda}}(W^{(K)}f)),$$

where $\widetilde{\lambda}$ is the threshold only determined by the noise level. Thus, the value of λ should be set larger than the value of $\widetilde{\lambda}$. It is empirically observed that $\lambda \approx 2\widetilde{\lambda}$ is a good choice. It is noted that the tight frame constructed during each iteration is directly estimated from the noisy data f instead of the noise-free image g.

The above adaptive tight frame denoising method is evaluated on several test images with different configurations. Through all experiments, all noisy images degraded by i.i.d. Gaussian white noise are synthesized as follows:

$$\mathbf{f} = \mathbf{g} + \epsilon(\sigma),$$

where $\epsilon(\sigma)$ is the i.i.d Gaussian noise with zero mean and standard deviation σ. Besides the visual inspection, the PSNR measurement is used to quantitatively evaluate the quality of the de-noised results. There are only two parameters λ and $\widetilde{\lambda}$ in the Algorithm 1 based adaptive tight frame denoising method. Both of them are closely related to the noise variance σ and the redundant degree of the used tight frame system. Through all experiments, we uniformly set $\lambda = 5.1\sigma$ and $\widetilde{\lambda} = 2.6\sigma$.

4.1. Computational cost

The proposed method is implemented in MATLAB 2011b and all the experiments are run on a laptop with a 2.1GHz Intel dual-core CPU and 4 GB memory. Table 1 listed the PSNR values of the results for the image "Barbara" of size 256×256 by Algorithm 1 with different maximum iteration numbers. The tight frame is initialized by 64 Harr wavelet filters. It is seen from Table 1 that after 25 iterations, there is little improvement on the PSNR values of the results. In other words, 25 iterations seem to be adequate to yield a good discrete tight frame for image denoising.

We compared the computational efficiency of Algorithm 1 against that of the K-SVD method [11] in terms of the running time on the same hardware configuration. There are two implementations of the K-SVD available online: one is implemented in MATLAB[1] and another is the faster version with key

[1] `http://www.cs.technion.ac.il/~elad/software/`

$\sigma \backslash K$	input	0	5	15	25	50	100	200	K-SVD
5	34.14	36.47	37.84	38.16	38.23	38.32	38.35	38.37	38.14
10	28.13	31.71	33.34	34.49	34.63	34.69	34.73	34.73	34.43
15	24.59	28.96	31.00	31.89	32.35	32.51	32.55	32.57	32.42
20	22.11	27.11	28.07	30.69	30.87	30.95	30.99	31.00	30.93
25	20.18	25.68	27.38	29.62	29.76	29.80	29.81	29.82	29.76

Table 1: The PSNR values (dB) of the noisy image, denoised results by Algorithm 1 with maximum iteration number $K = 0, 5, 15, 25, 50, 100, 200$ and denoised results by the K-SVD method with 15 iterations respectively, where σ denote the standard deviation of the image noise. The support of the filters in Algorithm 1 is set to 8×8. The patch size is set to 8×8 in the K-SVD method.

Method \ patch (filter) size	2×2	4×4	8×8	16×16
K-SVD in MATLAB	596.1	1930.99	10672.70	- - - -
K-SVD in C	29.35	20.81	57.13	576.88
Alg. 1 in MATLAB	**0.37**	**1.29**	**6.51**	**64.32**

Table 2: Comparison of running time in seconds between the K-SVD method with 15 iterations and Algorithm 1 with 25 iterations on the same hardware configuration.

components written in C language[2]. The running times of both implementations of the K-SVD method and the MATLAB implementation of Algorithm 1 are listed with respect to the image "Barbara" of size 256×256. It is seen that the methods with larger patch/filter size will require more time to execute the iterations.

It is also seen from Table 2 that our approach is much faster than the K-SVD method when the filter size is small. When the filter size is 8×8 or 16×16, the MATLAB implementation of our approach is still nearly 10 times faster than the C implementation of the K-SVD method with comparable PSNR values. The significant gain of our approach on computational efficiency comes from the fact that the dictionary we learned is a tight frame with perfect reconstruction property. There are only two simple components in our approach: one thresholding operation for denoising and one SVD operation for updating the tight frame. As the comparison, each iteration of the K-SVD method requires an iterative OMP method for each image patch (totally about 4000 patches for an image of size 256×256) to find the sparse approximation. It also updates all atoms in the dictionary in a sequential manner which is not fast either.

[2] `http://www.cs.technion.ac.il/~ronrubin/software.html`

initial tight frame	filter size	hard thresholding	Alg. 1
local DCT	2×2	26.90	26.32
	4×4	28.60	28.96
	8×8	30.08	30.47
	9×9	30.21	30.64
	16×16	30.45	30.93
Haar Wavelet	2×2	26.82	26.34
	4×4	27.59	29.91
	8×8	27.99	30.46
	16×16	28.05	30.95
linear framelet	3×3	27.52	28.14
	7×7	28.70	30.22
	15×15	28.95	30.86

Table 3: The comparison of the PSNR values (dB) of the de-noised results by the wavelet thresholding denoising method and by Algorithm 1 with respect to different initializations on tight frame and different filter sizes.

4.2. Performance of the proposed method with respect to different filter sizes and different initializations

In this experiment, we would like to see how the performance of the proposed method is influenced by the different settings, including the filter size and the initialization of tight frame. Three different types of initialized tight frames: shift-invariant local discrete cosine transform (DCT), shift-invariant Haar wavelet [1] and shift-invariant linear framelet [14], are tested on the image "Barbara" with different filter sizes. The Haar wavelet filters of size 2×2, 4×4, 8×8 and 16×16 are corresponding to the filters associated with the $1, 2, 3, 4$-level shift-invariant wavelet decompositions and reconstructions respectively. The linear tight frame filters of size 3×3, 7×7, 15×15 are corresponding to the filters associated with the $1, 2, 3$-level framelet decompositions and reconstructions respectively. Table 3 listed the PSNR values of the results from our approach after 25 iterations, using different initializations on tight frame of different filter sizes. It is seen from Table 3 that what tight frame is used for initialization does not have much impact on the performance. The PSNR values of the results by three different types of tight frames are nearly the same. However, the performance of our method varies significantly with respect to different filter sizes. It is seen that the larger filter size is used, or equivalently the more filters are used, the higher the PSNR values of the results will be. Such a phenomena is not surprised as the larger the filter size is and the more the filters are used, the learned tight frame filters are more likely to capture the special structures of the input image content. At last, the learned adaptive tight frame filters after 25 iterations are shown in Fig. 2 (a), compared with the initial inputed 4-level Haar wavelet filters shown in Fig. 2 (b).

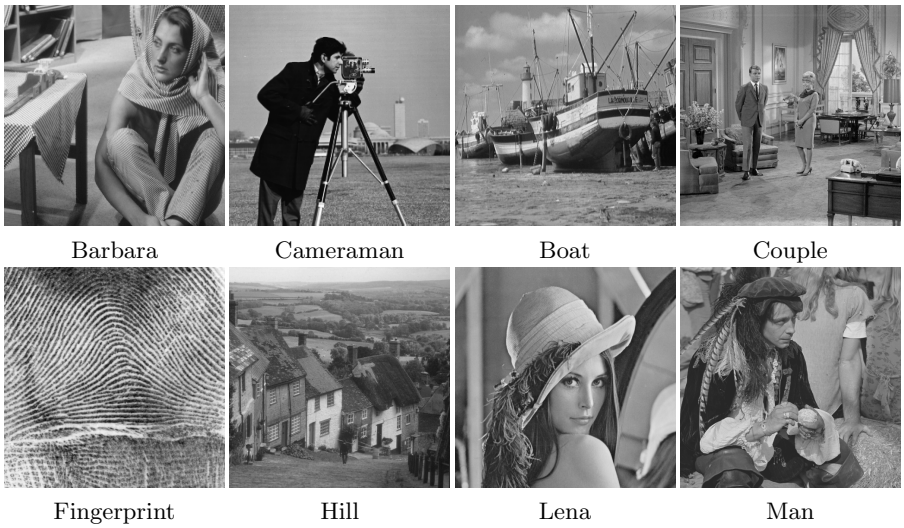


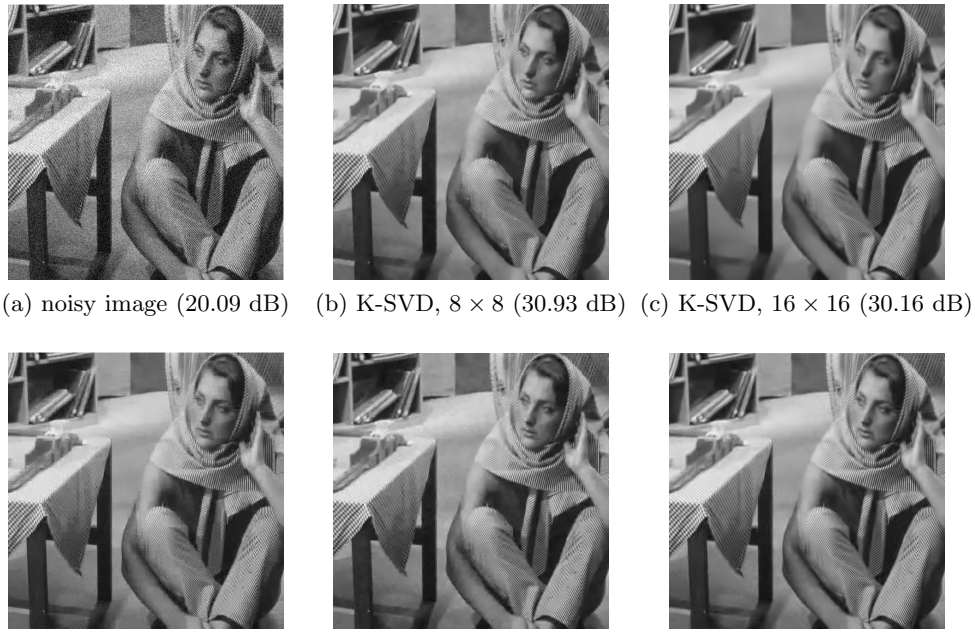
Figure 3: Visualization of 9 tested images.

4.3. Experiments on several sample images.

In this experiment, we tested the proposed method with different filter sizes on several images of different types, as shown in Fig. 3. The noisy inputs are synthesized by adding i.i.d. Gaussian white noise with different standard deviation σ to the original images. Two filter sizes are used in the experiments: 8×8 and 16×16. The results are compared against the results from the thresholding method by 3-level shift-invariant Haar wavelets, the K-SVD method with patch size 8×8 and the K-SVD method with patch size 16×16. Table 4 summarizes the PSNR values of the denoising results from these methods with different configurations.

As can be seen from Table 4, both the K-SVD method and our method performed better than the wavelet thresholding method on all images. In particular, there are significant improvements using the adaptive systems on the images including "Barbara", "Fingerprint" and "Lena". The main reason is that these three images all have some complex texture regions which the Haar wavelet transform cannot effectively sparsify.

On the other hard, the performances of the K-SVD method with patch size 8×8 and our approaches with filter 8×8, 16×16 are nearly the same with similar PSNR values. There are images on which the K-SVD method performed better and there are some on which our approaches performed better. Overall, the performances of our proposed method and the K-SVD method are comparable in terms of PSNR value, and so is the visual quality. See Fig. 4 for an visual inspection of the results for the image "Barbara" by different methods.



(a) noisy image (20.09 dB) (b) K-SVD, 8×8 (30.93 dB) (c) K-SVD, 16×16 (30.16 dB)

hard thresholding (27.98 dB) (d) Alg. 1, 8×8 (30.42 dB) (e) Alg. 1, 16×16 (31.01 dB)

Figure 4: Visual comparison of denoising results. (a) Noisy image; (b)–(c): the K-SVD method with patch size 8×8 and 16; (d): the result denoised by hard thresholding on 3-level shift-invariant Haar wavelet and (e)–(f): the results de-noised by Algorithm 1 with filter size 8×8 and 16×16.

image	σ	thresholding	K-SVD; 8×8	K-SVD; 16×16	Alg. 1; 8×8	Alg. 1; 16×16
Barbara	5	36.48	38.14	37.91	38.07	**38.26**
	10	32.10	34.43	33.96	34.26	**34.68**
	15	29.61	32.42	31.73	32.03	**32.51**
	20	27.98	30.93	30.16	30.42	**31.01**
	25	26.73	29.76	28.83	29.27	**29.85**
Cameraman	5	37.49	**37.93**	36.93	37.86	37.81
	10	32.97	**33.71**	32.79	33.59	33.54
	15	30.53	**31.46**	30.42	31.27	31.13
	20	28.89	**29.91**	28.92	29.59	29.61
	25	27.61	**28.91**	27.70	28.51	28.49
Boat	5	36.32	**37.16**	36.63	37.04	37.08
	10	32.81	33.63	32.96	33.65	**33.73**
	15	30.80	31.70	30.81	31.70	**31.77**
	20	29.34	30.31	29.27	30.32	**30.40**
	25	28.23	29.25	28.16	29.21	**29.34**
Couple	5	36.79	37.24	36.78	**37.31**	37.28
	10	33.08	33.50	32.74	33.63	**33.67**
	15	30.94	31.47	30.49	31.54	**31.63**
	20	29.43	30.02	28.97	30.07	**30.21**
	25	28.27	28.84	27.80	28.99	**29.15**
Fingerprint	5	35.01	**36.61**	36.06	36.58	36.55
	10	30.52	**32.39**	31.80	32.31	32.26
	15	28.12	**30.07**	29.35	29.91	29.92
	20	26.53	**28.44**	27.58	28.33	28.34
	25	25.35	**27.28**	26.32	27.17	27.17
Hill	5	36.33	36.96	36.51	36.96	**36.97**
	10	32.65	33.34	32.72	33.35	**33.35**
	15	30.74	31.43	30.68	31.51	**31.52**
	20	29.43	30.17	29.27	30.21	**30.25**
	25	28.41	29.19	28.24	29.23	**29.31**
Lena	5	37.63	38.56	38.13	38.61	**38.70**
	10	34.17	35.55	34.94	35.52	**35.71**
	15	32.17	33.72	32.98	33.61	**33.83**
	20	30.66	32.39	31.64	32.19	**32.43**
	25	29.46	31.35	30.45	31.05	**31.38**
Man	5	36.77	37.55	36.82	37.58	**37.57**
	10	32.73	33.60	32.68	33.63	**33.65**
	15	30.54	31.45	30.51	31.46	**31.49**
	20	29.09	30.13	29.09	**30.99**	30.08
	25	27.99	**29.11**	27.92	29.03	29.02

Table 4: Comparison of the PSNR values (dB) of Algorithm 1 and the K-SVD method with respect to different images and different noise level.

5. Conclusion and future work

Finding a sparse approximation of a given image plays an important role in many image restoration tasks. Wavelet tight frames have been successfully used to restore the image of interest by utilizing its sparsity under the wavelet tight frame, e.g, the framelets and shift-invariant wavelets. However, due to the significant variations of image structure, a pre-defined redundant system is not efficient when representing images of complex structures. In this paper, we developed an iterative numerical scheme to construct a discrete tight frame that is adaptive to the given image. Different from most existing learning based approaches which learn an over-complete dictionary, the dictionary constructed in our approach is always a tight frame during each iteration. As a result, the proposed approach is very computationally efficient. Also, the tight frame we constructed has several properties appealing to many image processing applications than the over-complete dictionary, including the perfect reconstruction property. Based on the proposed construction scheme, the derived adaptive tight frame denoising method shows its advantage over the traditional wavelet thresholding method. Also, our tight frame based approach is much faster than the over-complete dictionary based approaches, e.g., the K-SVD method. In future, we would like to investigate how to construct the adaptive tight frame from images corrupted by other factors, e.g., the blurring.

References

[1] I. Daubechies, Ten lectures on wavelets, 1st Edition, CBMS-NSF Lecture Notes, SIAM, 1992.

[2] A. Ron, Z. Shen, Affine system in $L_2(R^d)$: the analysis of the analysis operator, J. of Func. Anal. 148.

[3] I. Daubechies, B. Han, A. Ron, Z. Shen, Framelets: MRA-based constructions of wavelet frames, Appl. Comput. Harmon. Anal. 14 (2003) 1–46.

[4] E. Candes, Ridgelets: estimating with ridge functions, Ann. Statist. 31 (1999) 1561–1599.

[5] E. Candes, D. L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise-C^2 singularities, Comm. Pure Appl. Math 57 (2002) 219–266.

[6] E. J. Candes, D. L. Donoho, Recovering edges in ill-posed inverse problems: Optimality of curvelet frames, Ann. Statist. 30 (3) (2002) 784–842.

[7] S. Mallat, E. Lepennec, Sparse geometric image representation with bandelets, IEEE Transactions on Image Processing 14 (2005) 423–438.

[8] G. Kutyniok, D. Labate, Construction of regular and irregular shearlet frames, J. Wavelet Theory and Appl 1 (2007) 1–10.

[9] M. Lewicki, J. Sejnowski, Learning overcomplete representation, Neural Computation 12 (2000) 227–365.

[10] K. Kreutz-Delgado, J. Murray, B. Rao, K. Engan, T. Lee, T. Sejnowski, Dictionary learning algorithms for sparse representation, Neural Computation 15 (2) (2003) 349–296.

[11] M. Elad, M. Ahron, Image denoising via sparse and redundant representations over learned distionaries, IEEE Trans. on Image Processing 54 (12) (2006) 3736–3745.

[12] J. Mairal, M. Elad, G. Sapiro, Sparse representation for color image restoration, IEEE Transactions on Image Processing 17 (1) (2008) 53–69.

[13] J. Mairal, F. Bach, J. Ponce, G. Sapiro, A. Zisserman, Non-local sparse models for image restoration, in: ICCV, 2009.

[14] Z. Shen, Wavelet frames and image restorations, Proceedings of the International Congress of Mathematicians, Hyderabad, India.

[15] B. Dong, Z. Shen, MRA based wavelet frames and applications, IAS Lecture Notes Series, Summer Program on "The Mathematics of Image Processing", Park City Mathematics Institute.

[16] A. Chai, Z. Shen, Deconvolution: A wavelet frame approach, Numer. Math. 106 (2007) 529–587.

[17] J.-F. Cai, R. Chan, Z. Shen, A framelet-based image inpainting algorithm, Appl. Comput. Harmon. Anal. 24 (2008) 131–149.

[18] J.-F. Cai, H. Ji, C. Liu, Z. Shen, Blind motion deblurring from a single image using sparse approximation, in: CVPR, 2009.

[19] B. Han, G. Kutyniko, Z. Shen, Adaptive multiresolution analysis structures and shearlet systems, SIAM Journal on Numerical Analysis 49 (5) (2011) 921–1946.

[20] B. Olshausen, B. Field, Sparse coding with an overcomplete basis set: a strategy employed by v1, Vision Research 37 (1997) 3311–3325.

[21] D. Donoho, De-noising by soft thresholding, IEEE Trans. Inf. Theory 41 (3) (1995) 613–627.

[22] J.-L. Starck, M. Elad, D. L. Donoho, Image decomposition via the combination of sparse representations and a variational approach, IEEE Trans. Image Proc. 14 (2005) 1570–1582.

[23] I. Daubechies, M. Defrise, C. D. Mol, An iterativethresholding algorithm for linear inverse problems with asparsity constraint, Comm. Pure Appl. Math 57 (11) (2004) 1413–1457.

[24] M. Fadili, J.-L. Starck, F. Murtagh, Inpainting and zooming using sparse representations, The Computer Journal 52 (2009) 64–79.

[25] R. H. Chan, T. F. Chan, L. Shen, Z. Shen, Wavelet algorithms for high-resolution image reconstruction, SIAM J. Sci. Comput. 24 (2003) 1408–1432.

[26] L. Borup, R. Gribonval, M. Nielsen, Bi-framelet systems with few vanishing moments characterize besov spaces, Appl. Comput. Harmon. Anal. 14 (1).

[27] J.-F. Cai, B. Dong, Z. Shen, S. Osher, Image restoration: total variation; wavelet frames; and beyond, J. Amer. Math. Soc. to appear.

[28] H. Zou, T. Hastie, R. Tibshirani, Sparse principle component analysis, Journal of Computational and Graphical Statistics 15 (2) (2006) 265 – 286.